Imagining Better Education:

Proceedings of the 2019 IBE Conference

Editors: Harriet Axbey and Sharon Riddle
Editorial board: Katie Allen, Sophie Anderson, Martina Diehl
Editorial advisors: Prof. Prue Holmes & Dr Dimitra Kokotsaki
Cover design: Sharon Riddle
Cover photographer: Image by Emphyrio from Pixabay

Imagining Better Education: Proceedings of the 2019 IBE Conference
Edited by Harriet Axbey & Sharon Riddle
ISBN (print) 978-0-907552-24-6
ISBN (digital) 978-0-907552-23-9

First published 2020
Published by Durham University

Durham: IBE Conference * School of Education * Durham University Leazes Road * Durham DH1 1TA * United Kingdom

Contents

IBE conference abstract – Dimitra Kokotsaki

The *Imagining Better Education* conference, an event organised by our Post-Graduate Research students at the School of Education at Durham University, has provided a fruitful space for collaboration, learning and engagement for our student research community. This was the second student-led conference, following its successful launch one year ago. The two-day conference was an excellent networking opportunity where the students came together, presented their research findings, learnt from one another and engaged in constructive dialogue and fruitful debate. There were 22 presentations produced by scholars and practitioners from across a range of different countries. A workshop on networking and communication skills in addition to a keynote speech on evidence-based education were also useful highlights of the conference. Nine papers are included in this proceedings volume. The research areas are wide ranging and employ diverse methodologies, appropriate for the unique educational and professional settings that are investigated. The papers included in the proceedings explore systematically the following themes: cross-cultural teaching, internationalisation and English for Academic Purposes in higher education, teacher effectiveness, reading attainment, the role of written feedback in promoting student self-regulation, language teaching and academic writing in a variety of contexts. This breadth of coverage attests to our students' passionate engagement with areas of personal and professional interest. It also reflects the School of Education's standing as one of the major leaders in educational research. The publication process for the proceedings aimed to offer space for the students' academic development. Two peer reviewers who were part of the editorial team blind reviewed all submitted manuscripts. The opportunity to offer and receive feedback provided a learning opportunity for both the authors and the editorial team. We hope that all participants found the experience of presenting, writing, reviewing and engaging in dialogue and networking with like-minded researchers exciting, motivating and constructive. We are all extremely proud of our very talented and competent research students that made this conference a success for another year. Their passion and commitment to imagine and help realise a better education for all is testimony to their insatiable drive and motivation.

Author Information

Ismail Aslantas

Ismail is a PhD student at the School of Education, Durham University. His research is on the evaluation of teachers' performance based on student test scores using Value-added Models (VAMs). Using longitudinal quantitative data spanning three successive years, the PhD project examines the stability of the teacher value-added effectiveness estimates regarding the number of test scores used, the predictors used in the estimations, and the analysis methods applied.

Harriet Axbey

Harriet Axbey is an ESRC-funded student in the School of Education, having previously qualified as a teacher. Working within a neurodiversity paradigm, her main research aims to look at interactions between autistic and non-autistic children in the classroom. Additionally, she has conducted research while in China, where she is interested in the 'Teaching for Mastery' method of teaching mathematics, and the implications of implementing it in the UK.

Xuanhong Guo

Xuanhong Guo was a CSC funded PhD student in School of Education, Durham University. She graduated in January 2020. Her research focused on the use of lexical cohesion in EFL learners' English academic writing and EAP pedagogical implications. Her supervisors were Dr Philip Nathan and Dr David Stevens.

Honghuan Li

Honghuan Li is an Ed.D. student in School of Education, Durham University, with a research focus on teacher ICT beliefs, attitudes and training needs on virtual reality technology.

Nuria Polo-Pérez

Nuria Polo-Pérez is an Assistant Teaching Fellow in Spanish at the Centre for Foreign Language Study, and a part-time doctoral researcher in the School of Education at Durham University. Her research project is an ethnographic study of two 'language cafés', and focuses particularly on the affordances of these informal environments for meaningful multilingual socialisation, and the development and performance of language learners' multilingual identities.

Sharon Riddle

Sharon Riddle is a doctoral student at the School of Education, Durham University. Her research interests include English for Academic Purposes, and where it is positioned in UK higher education.

Lindsey Wardle

Lindsey is an ESRC-funded doctoral researcher at Durham University and a member of Durham University Evidence Centre for Education. Her research is focused on reading

attainment in North East England, and understanding why these trends are different compared to other regions.

Meechai Wongdaeng

Meechai Wongdaeng is an Ed.D. student at the School of Education, Durham University under the sponsorship of the Royal Thai Government. His thesis project is about investigating and evaluating the effectiveness of metacognitive instruction on the learning outcomes of tertiary students in an EFL context. He is keen to gain insight into the EFL policy gaps and evidence-based practices.

Suleyman Yildirim

Suleyman is a third year PhD student in the School of Education at Durham University. The aim of his research is to examine the effect of written feedback on student self-regulated learning and critical thinking ability in the higher education in the UK. Moreover, he is interested in student intrinsic motivation and educational assessment.

About the Imagining Better Education conference

In July 2019 we welcomed postgraduate researchers from the School of Education at Durham University to the second annual student conference. The conference aimed to showcase the exciting work being done by postgraduate research students in the School, as well as offering an opportunity for networking. A common aim for researchers is to strive towards improvements in whatever division of education one chooses to focus their attention on; hence the title of the conference was chosen: Imagining Better Education. Sessions were organised to strike a careful balance between students presenting their own work and listening to experienced academics and professionals talk about their experiences on topics such as effective networking and using research evidence to inform classroom practice. We hope the students found the experience as useful and rewarding as we did.

Katie Allen

Introduction

Durham University's second annual Postgraduate Educational Research Conference, entitled *Imagining Better Education*, marked an exciting second milestone for the Postgraduate community here at the School of Education. The student-led conference saw PGRs from across the Department come together to organise a two-day event during which students presented their research findings to staff and colleagues. There were 22 presentations produced by scholars and practitioners from a range of different countries and areas within Education. Nine papers are included in this proceedings volume.

The conference provided a supportive and collaborative environment for PGRs to discuss their research with their peers. This acted as a forum for students to learn from each other whilst sharing their own insights into what it means to design and carry out research. Similarly, the publication process for these proceedings was intended to support students' academic growth. Every presenter was given the opportunity to submit a manuscript for the conference proceedings. Through a blind review process, each author received feedback from two peer-reviewers. This was supported by a two-hour workshop on writing and reviewing for publication. Our aim was to make the review and editing process a learning opportunity for authors by working with an editorial team. We hope that students benefitted from the experiences of presenting, writing, and reviewing; making them more confident and accomplished authors.

It is important to acknowledge the dedication of our PGR students, not only to their studies at the School of Education, but also to their desire to improve Education. The contributions by the authors in the following proceedings, and from all who presented at the conference, reflect their commitment to "Imagining Better Education" for learners. Major themes of the featured papers include intercultural education, attainment, and teacher effectiveness. The breadth of coverage reflects he School of Education's standing as one of the major leaders in educational research, with the efforts of our PGR community continuing in this endeavour by shaping policy and practice for the future. In this way, these proceedings build a legacy of scholarly contribution for the School of Education and its PGR students. I would like to express my gratitude to all of the authors in contributing their papers for publication in the *Imagining Better Education 2019 Conference Proceedings*. They provide a fascinating insight into how we can imagine, and strive for, better education for all.

Katie Allen

Imagining Better Education Conference Proceedings 2019

The Stability Problem of Value-added Models in Teacher Effectiveness Estimations: A Systematic Review Study

Ismail Aslantas

School of Education, Durham University, Durham, United Kingdom

ismail.aslantas@durham.ac.uk

The Stability Problem of Value-added Models in Teacher Effectiveness Estimations: A Systematic Review Study

This article provides evidence by undertaking a systematic review on the stability problem of using value-added models in teacher effectiveness estimates from the perspective of the impact of the number of previous test scores employed aimed at answering a unique review question (This is a long sentence that lacks clarity – it could be broken down for a clearer meaning): How stable is teacher effectiveness estimates measured by VAMs? By using the terms: teacher performance, student performance, value-added model, stability and their other related synonyms, a comprehensive search was conducted in 17 databases along with employing hand search in Google Scholar and contacting authorised persons by email. In total 1439 records were found as a result of the searches. After completing the screening process, 50 studies remained for data extraction.

Out of 50 a total of studies in the review list, 13 focused on the stability of VAM estimates regarding using the number of prior test scores. In summary, there is a common view that the use of prior year data in on value-added estimates for teacher effectiveness has a positive impact, however, with regard to the impact of multiple previous year data, different voices arose from the researchers.

Keywords: teacher effectiveness, value-added model, stability, systematic review

Introduction

Although the use of value-added models (VAMs) in the field of business and economics at first, as researchers in the education sector began to be interested in these models to measure teacher performance, it appears that there were research (studies) on VAMs in this field in the late 1900s. (Sanders & Horn, 1994, 1998; Sanders, Saxton, & Horn, 1997; Webster & Mendro, 1997; Kupermintz, 2003). With the effect of No Child Left Behind Act of 2001 (NCLB) based on the belief that the most important school-related factor affecting student achievement is the quality of the teacher (Aaronson, Barrow, & Sander, 2007; Rivkin, Hanushek, & Kain, 2005), the studies done in this field gained momentum (Ballou, Sanders, & Wright, 2004; McCaffrey, Lockwood, Koretz, Louris, & Hamilton, 2004; Newton, Darling-Hammond, Haertel, & Thomas, 2010). In order to measure the effectiveness of teachers, several types of VAMs were developed and applied by states and school districts in the US such as the Tennessee value-added assessment system (TVAAS) and the Dallas value-added accountability system (DVAAS). The common assumption of these models that the growth/decline of the students' achievement in the standardised tests is attributed to their teacher performance. In all VAMs concepts, it is assumed that the quality of students' performance in the school reflects the quality of the teaching received (Darling-Hammond, 2015). A teacher's value-added scores can be calculated by subtracting his/her students' predicted scores from their actual scores in the standardised test(s) (Sanders, Saxton, & Horn, 1997). The predicted score is subjected to the same students' one or more previous year test scores and their characteristic features in most VAMs (Ouma, 2014). These prediction methods are likely to be beneficial to identify the most and least effective teachers in a school, district and/or state (Colorado, 2007).

As a result of the increasing number of research studies in parallel with the interest of researchers in this field, it was possible to discover the strengths and weaknesses of VAMs. The teacher performance results estimated by VAMs are highly affected by students' characteristics and other conceptual predictors which teacher cannot control (Wei, Hembry, Murphy, & McBride, 2012). These models should not be used for the high-stake decision about the teachers unless the pros and cons of VAMs are fully revealed. Therefore, their use should be limited to improve the educational institutions, provide teachers with the opportunity to address their own shortcomings and provide justification for students' academic progress. More accurate information is needed about value-added models to expand the intended use of them. One of the robust methods to collect accurate information from the literature is conducting a systematic review. Primarily the research has the aim to determine whether the teacher performance results estimated by applying value-added models under different conditions are stable or not. Specifically, observable teacher characteristics, school characteristics, the students' test scores obtained over time and the preferred data analysis methods are the different conditions that are referred to in this current research. As one of the purposes of this study is to provide guidance and appraise to policymakers and practitioners on the use of value-added models in the teacher performance appraisal for high-stake purposes such as decisions on dismissal and monetary reward, gathering evidence-based findings from a wide body of research systematically are needed instead of from narrative literature. Therefore, this systematic review is utilised in order to

synthesise the results of previous relevant studies that analysed the impact of conceptual predictors and data analysis methods used on teacher performance evaluation. In line with the purpose of this systematic review declared, a unique review question formulated to assist in exhaustively examining the available evidence is;

How stable is teacher effectiveness estimates measured by VAMs?

where teacher effectiveness operationally defined by VAM as the estimation of the differences between expected and observed student test scores (Kersting, Chen and Stigler, 2013). Moreover, in this systematic review study, the operational definitions of the term stability refer to the stableness of the estimates due to (a) *the number of test scores used*, (b) *the predictors used in the estimations*, and (c) *the analysis methods applied*. The existing literature on the stability of VAMs estimates will be retrieved from these three perspectives.

Methods Databases and Searching Strategy

In order to conduct a comprehensive search in the impact of conceptual predictors and data analysis methods preferred on teacher performance evaluation estimates, both published and unpublished studies that met the inclusion criteria as explained below, are obtained until the 1st of May 2019. In order to identify the studies met the inclusion criteria, a total of 17 electronic databases and six their providers were employed (shown in Table 1). For reaching other relevant sources, research centres, foundations, and researchers who have worked on teacher performance evaluation based on VAMs were contacted personally.

Table 1. *Databases and their providers*

	Provider	Database
1	ProQuest	ProQuest Dissertations & Theses Global: Social Sciences
		Education Database
		ERIC
		International Bibliography of the Social Sciences (IBSS)
		Social Science Database
		Applied Social Sciences Index & Abstracts (ASSIA)
2	EBSCOhost	OpenDissertations
		British Education Index
		Business Source Premier
		Education Abstracts
		Educational Administration abstracts
		PsycINFO
3	Web of Science	Web of Science Core Collection
		Current Contents Connect
4	Elsevier	SCOPUS
5	SAGE Research Methods Core	SAGE Journals
6	Taylor & Francis Online	Educational Research Abstracts Online

To formulate the searching strings, first of all, the keywords which are “teacher performance”,

"student performance", "value-added model" and "stability" were identified in parallel with the review question. Then the related terms of the keywords were determined by identifying which alternative terms were used to substitute the search terms in the related sample studies found (shown in Table 2).

Table 2. *Search keywords*

Search Terms	**Related Terms**	
Teacher performance	Teacher effect*	Teacher proficiency-rank
	Teacher evaluation	Teacher judgment
	Teacher performance evaluation	Educational effectiveness
	Teacher appraisal	Educator performance appraisal
	Teacher performance appraisal	Educator performance
	Teacher quality	Educator evaluation
	Teacher assessment	Educator quality
	Teacher performance assessment	Teaching effect*
	Teacher accountability	Measuring teach*
	Teacher proficiency	Evaluating teach*
Student performance	Academic achievement	Achievement
	Academic gains	Achievement measures
	Student test score	Outcomes
	Student test score	Outcome measures
	Student test-score	Student test performance
Value Added Model	Value added modelling	VAM*
	Value-added model*	Value added estimate*
	Teacher value-added	Value-added estimat*
Stability	Concord*	Imprecision
	Robust	Variat*
	Sensitivity	Fluctua*
	Instability	Persistence
	Precision	Shrink*

Inclusion and Exclusion Criteria

The research included in this systematic review met all criteria located in Table 3.

Table 3. *Inclusion criteria*

Inclusion Criteria	
Criteria	**Description**
The population of this study is teachers evaluated by their students' outcomes	Only studies focused on teacher performance evaluation based on student test scores will be included in this systematic review. Among the studies that are interested in the performance evaluation on more than one subject such as teacher and school performance estimated together in a single study, only the studies whose one of the areas of interest is teachers will be given a place in this systematic review.
The issue of the study is the stability of the estimates	The operational definitions of the term of stability in this systematic review refer to the stability of the estimates; (a) based on students' test scores Studies that use this measure for stability must have at least two previous years test scores for the same or different cohorts of students over time) (b) based on predictors used (c) based on the analysis methods applied Studies are included if they use any one of the above measures of stability.
Only empirical studies are reviewed for this study	Empirical studies refer to primary research as opposed to secondary research such as reviews, government reports. Studies analyzing secondary data such as panel, administrative data are considered as primary research.
The study setting of the research interest is K-12.	(K) refers to kindergarten grade (age 5-6, equivalent to Year 1 in the UK) and (12) refers to the 12th grade (age 17-18, equivalent Y13 or 6th form in the UK. All studies conducted from kindergarten to the 12thgrade setting are included in this systematic review.
Published in English	Studies reported in English

The studies were screened with regard to the below criteria as the exclusion criteria of this research;

Not reported or published in English

Not primary research

Not about education

Not within K-12 (e.g. higher education, reception year or nursery)

Not about the evaluation of teacher effectiveness

About the use of value-added measures of teachers to predict teacher attrition

The outcome is not student test scores or gains (e.g. children's behavior or attendance)

Using measures of teacher effectiveness to predict outcomes

Just about school effectiveness or school improvement (the studies focused on both school (principal) and teacher effectiveness, they have potential to be included in the review)

About teacher effectiveness in non-mainstream school

Just about pupils with special educational needs (SEN)

About theories and policies, opinion pieces, discussion pieces

Instructional manual or promotional literature about how to measure teacher effectiveness

Literature about the characteristics of effective teachers

Findings

The research findings from this systematic review are presented mainly in the two sections: description results and thematic analysis results. In the first section, overall explanatory results for the searching process were exhibited. On the basis of the conceptual definition of the stability of VAMs estimates, three themes have been determined for this study; a) the number of previous test scores employed, b) the predictors used, and c) the data analysis methods applied. As this article is a part of the uncompleted doctorate thesis, the key findings of the included studies were only placed under the first theme; the number of previous test scores employed.

Descriptive Results

For a more efficient and well-organized search process, the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) flow diagram is mainly guided. PRISMA is a diagram that published in 2009 by the PRISMA group (Moher, Liberati, Tetzlaff, & Altman, 2009) to help researchers to map out the number of studies identified, included and excluded based on the criteria established. The number of studies included and excluded in the review list was illustrated in the flowchart in Figure 1.

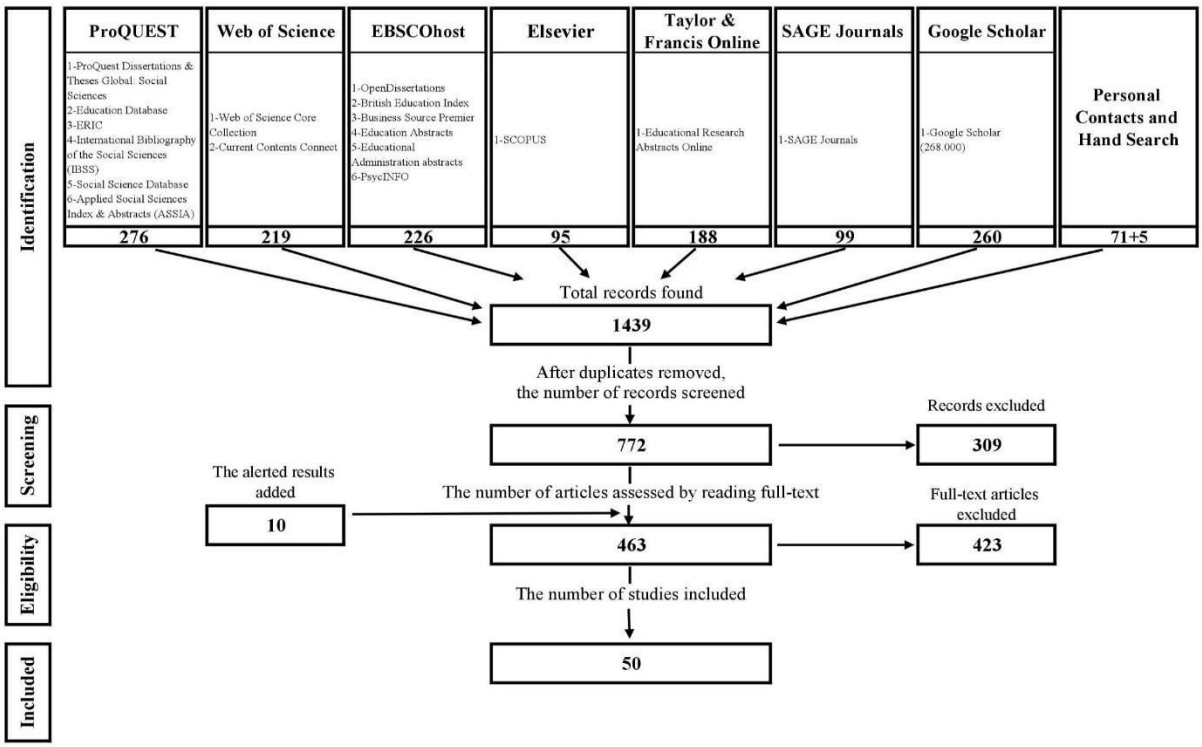


Figure 1. *The Modified PRISMA Flow Diagram*

As a result of typing a combination of the searching strings into the providers' search engines, a total number of 1439 articles were found initially. In the process of merging, 667 cases were deleted from the review list as being duplicated cases of unique studies. Through the phase I screening checklist form, all remaining 772 studies' titles and abstracts were screened, and 309 of them were excluded from this systematic review study in this screening stage. The further 423 cases out of the remaining 473 studies, which were added 10 studies came from the alerted results on the providers, were removed by employing the phase II full-text screening checklist form. To ensure that the screening processes were undertaken away from the prejudice of the researcher and to minimise the lack of potentially relevant articles among the discarded papers, the randomly selected 70 papers were also screened by a second independent reviewer. Although over 95 percent exact agreement reached between the reviewers, the inter-rater reliability agreement was .86.

Finally, the remaining studies were also subjected to the quality appraisal in order to include the findings from only the credible studies in this systematic review. Unfortunately, the trustworthiness of all research done is not the same, so the credibility of their findings should not be the same either. Therefore, to avoid an invalid and misleading conclusion, as of being a practical way for evaluation of the quality of individual studies instead of complex technical checklists in the literature, the modified "sieve" quality appraisal framework designed by Gorard (2014) was employed in this systematic review. As the qualifications of the studies were satisfactory for inclusion in this study, none of them was excluded from the review list. Therefore, by the end of this searching and elimination strategies, 50 research that focused on the stability of teacher effectiveness measurement estimates by VAMs were included in this current study.

Thematic Analysis Results

The final 50 studies retrieved in this systematic review by three main themes; the number of previous year test scores employed, the prediction used, and the data analysis methods applied in estimating teacher effectiveness. Although it was planned to present the results under three themes, as the analysis process in other themes is not yet complete, only the key findings related to the fist theme, the number of previous test scores employed, were placed.

The Number of Previous Test Scores Employed

In this section, the stability of teacher performance evaluation estimations with regard to the number of students' previous test scores used in VAMs was investigated. Although out of 50 studies, a total of 15 research studies are located in this theme.

The first study (Rothstein, 2009) claimed that the limited positive effect of using additional scores was longitudinal research involved 49.456 students from grade 3 to 5 linked to 2844 teachers. The researcher discussed a bias problem in value-added estimates related to student-teacher allocation issues. Although the study covers other aspects of VAM estimates, in this systematic review, only findings related to the impact of a number of test scores from prior years were included. To investigate the impact of using additional year test scores on teacher effectiveness in grade 5, first, the authors added the nearest prior year score (test score in 4^{th} grade) and estimated the increase in R^2, which was .55. The 4^{th}-grade test score contributed to 55 percent to the explanation of variance on test scores in grade 5. Moreover, the authors also included two prior year test scores (pre- and post-tests in grade 3) in the previous model; however, their contribution was very limited, R^2 only increased by .039 (3.9%).

The limited effect reported by the other study (Cunningham, 20014) was a thesis defended for the Doctor of Philosophy degree. In order to provide sufficient information to policymakers and practitioners in making the high-stakes decision regarding with teacher accountability system, this study evaluated the impact of growth model preference, of how many years of data used, and of student-level variable employed in the teacher performance estimates. With using up to three successive years of student data, teacher effects were estimated from five value-added models. Teacher rank orderings obtained in the five value-added models by using either one or three years of test scores were highly correlated with each other. The correlation exceeded .90 when using single-year data, and .80 when conducting multiple years of data. The use of a single year of test scores instead of three years resulted in a slight increase in correlation between the models and a slight decrease in teacher movement between quarters.

The next study, Shafer et al. (2012), compared six growth models used for teacher effectiveness estimates in the literature; quantile regression (QReg), ordinary least square (OLS), growth score difference (year two minus year one), and three different transition models (value-tables). Although the study compares the six growth models, it also covers the impact of using additional previous year data in estimates of the growth models. With related to this theme's concept, the findings on the correlations between scores estimated in reading and maths across four student cohorts allocated were presented. Mainly, the study claimed that the inclusion of data from more previous years (at least two years) in the QReg and OLS

models had a limited positive effect. For instance, the correlation between maths and reading scores for cohort 1 students was found as .19 in QReg1 (used one prior year only) and .18 in QReg2 (used two prior years); similarly the greatest change in correlation coefficients with regard to use additional prior year data in OLS models was found as .01.

Oppositely, eight research in the review list claimed the advantages of using additional prior years data in the teacher effectiveness estimates. One of them done by Lash et al. (2016) in order to investigate how stable the teacher growth percentile scores over the years, so that the authors compared the reliability of coefficients of the estimations. They claimed that stability of performance scores increased from .5 to .67 when the results are obtained by averaging over two years, and to .75 by averaging three years in maths, similarly in reading the increase obtained from .41 to .58 by averaging two years, and .68 by averaging three years.

The next study claimed a substantial improvement obtained in the reliability of VAM estimates by employing additional years of observations is belonging to Goldhaber and Hansen (2013). In order to examine the long-term stability of teacher effectiveness, the authors used up to ten years of longitudinal data in maths and reading across 3rd to 5th grade. After running a series of value-added estimates, the authors reached that there is a substantial improvement in the reliability of the estimates by using multiple prior year test scores. The reliability coefficient increased from 0.29 with a one-year VAM to 0.52 with a six-year VAM.

Another longitudinal study (Hu, 2015) involved 1210 maths, and 1239 reading teachers were conducted to explore the impact of adding student prior achievement into the estimates of teacher effectiveness by using the longitudinal students' data (one to three previous year test scores depend on grade and year) into hierarchical linear models. As a result of this study, students' up to three prior years test scores explained more than half of variance in their current scores was found. Besides, the researcher claimed that not surprisingly, the nearest previous year test score had an important role in this explanation. The average of 57% and 59% of the variance in students' current achievements in maths and reading, respectively, were explained by the nearest prior test scores. Similarly, the additional previous year test scores also contributed to the explanation of the variance in students' current achievement, but not as large as the nearest prior year's. For instance, 67% of total variance in students' mathematics achievement in Grade 7 in 2009-10 were explained by their achievement in grade 6 and 5 (the impact of the achievement score in grade 5 was 12%), and 69% of variance in grade7 in 201011 was explained by the achievement scores in grade 6, 5 and 4 (the impact of achievement score in grade 4 was 2%).

Discussion and Initial Conclusion

This systematic review study is utilised as a part of the researcher's own doctorate thesis in order to synthesise the results of previous relevant studies that analysed the impact of conceptual predictors and data analysis methods used on teacher performance evaluation In line with the purpose of this systematic review, a unique question was formulated in order to retrieve the available evidence. Namely, how stable is teacher effectiveness measured by VAMs? In this systematic review study, the operational definitions of the term stability refer

to the stableness of the estimates due to (a) the number of test scores used, (b) the predictors used in the estimations, and (c) the analysis methods applied. The existing literature on the stability of VAMs estimates will be retrieved from these three perspectives.

Since the researcher's doctoral process is ongoing and the analysis chapter is not completely finished, yet, only the key findings related to the first theme, *the number of previous test scores employed*, were placed in this conference proceedings paper. The stability of teacher performance evaluation estimations with regard to the number of students' previous test scores used in VAMs was investigated. Out of 50 studies, a total of 15 research studies are retrieved in this theme. In general, although there is a consortium on the impact of prior year data on value-added estimates for teacher effectiveness, unfortunately, this consortium is disintegrated about the impact of using additional previous year(s) data on the estimates. Keep in mind that the evidence in this theme is not very robust because of preferring not strong design for their research questions and involving a considerable amount of missing data. Eight studies in this theme, seven of them were rated 2 * from middle bound, and one with 1 * from the lower bound, claimed to be advantageous with adding additional year test scores to the estimates. Although the other seven studies, one of these was rated 2 * from upper bound, and the rest were rated with 2 * from middle bound reported that using additional prior year test scores have a positive impact, but the research also found that the impact is limited, or even little if any. Therefore, the findings are mixed with almost an equal number of medium quality studies suggesting that there are advantages in including additional year test scores as well as those advocating having little benefit. However, the stronger study (rated 2a) suggests that there is little benefit of using additional test scores from previous years. More robust studies may be needed to confirm the results, but at the moment, there is no evidence that using additional prior test scores is useful.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teacher and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135.

Ballou, D., Sanders, W. L., & Wright, P. (2004). "Controlling for Students' Background in value-added assessment for teachers. *Journal of Educational and Behavioral Statistics, 29*(1), 37-65.

Colorado, G. (2007). *An empirical sensitivity analysis of value-added teachers' effect estimates to hierarchical linear model parameterizations.* University of Northern Colorado.

Cunningham, P. L. (2014). *The effects of value-added modeling decisions on estimates of teacher effectiveness.* The University of Iowa.

Darling-Hammond, L. (2015). Can Value Added Add Value to Teacher Evaluation? *Educational Researcher, 44*(2), 132-137.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. J. (2014). The sensitivity of value-added estimates to specification adjustments: evidence from school- and teacher-level models in Missouri. *Statistics and Public Policy, 1*(1), 19-27.

Goldhaber, D., & Hansen , M. (2010). *Assessing the Potential of Using Value-Added Estimates of Teacher Job Performance for Making Tenure Decisions.* The Urban Institute. 2100 M Street NW, Washington, DC 20037: National Center for Analysis of Longitudinal Data in Education Research.

Goldhaber, D., & Hansen, M. (2013). Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica, 80*(319), 589-612.

Gorard, S. (2014). A proposal for judging the trustworthiness of research findings. *Radical Statistics, 110*, 47-59.

Hu, J. (2015). *Teacher evaluation based on an aspect of classroom practice and on student achievement: A relational analysis between student learning objectives and value added modeling.* Boston College.

Johnson, M. T., Lipscomb, S., & Gill, B. (2015). Sensitivity of Teacher Value-Added Estimates to Student and Peer Control Variables. *Journal of Research on Educational Effectiveness, 8*(1), 60-83.

Kersting, N., Chen, M.-k., & Stigler, J. (2013). Value-added teacher estimates as part of teacher evaluations: exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives, 21*(7).

Koedel, C., & Betts, J. R. (2011). Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein Critique. *Education Finance and Policy, 6*(1), 18-42.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis, 25*(3), 287–298.

Lash, A., Makkonen, R., Tran, L., & Huang, M. (2016). *Analysis of the stability of teacher-level growth scores from the student growth percentile model* (REL 2016–104) . Institute of Education Sciences, Department of Education. Washington, DC: USA: National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.

McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (n.d.). The intertemporal variability of teacher effect estimates. *Education Finance and Policy, 4*(4), 572606.

McCaffrey, D., Lockwood, J., Koretz, D., Louris, T., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*, 647-101.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Med, 7*.

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18*(23), 1-27.

Ouma, C. A. (2014). *Performance of cart-based value-added model against HLM, multiple regression, and student growth percentile value-added models.* PhD Thesis, Florida State University.

Potamites , L., Booker , K., Chaplin , D., & Isenberg , E. (2009). *Measuring school and teacher effectiveness in the EPIC Charter School Consortium-Year 2.* Washington, DC: USA: Mathematica Policy Research.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rothstein, J. (2009). Student sorting and bias in value added estimation: Selection on observables and unobservables. *Education Finance and Policy, 4*(4), 537-571.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8,* 299-311.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Sanders, W. L., Saxton, A., & Horn, B. (1997). The Tennessee value-added assessment system: A quantitative outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: is student achievement a valid evaluation measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press, Inc.

Schafer, W. D., Lissitz, R. W., Zhu, X., Hou, X., & Li, Y. (2012). Evaluating teachers and schools using student growth models. *Practical Assessment, Research & Evaluation, 7*(17).

Stacy, B., Guarino, C., & Wooldridge, J. (2018). Does the precision and stability of value-added estimates of teacher performance depend on the types of students they serve? *Economics of Education Review, 64*, 50-74.

Webster, W., & Mendro, R. (1997). The Dallas value-added accountability system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81-99). Thousand Oaks, CA: Corwin Press, Inc.

Wei, H., Hembry, T., Murphy, D. L., & McBride, Y. (2012). *Value-added models in the evaluation of teacher effectiveness: A comparison of models and outcomes (Pearson Research Report).* New York, NY: Pearson.

Fusion or Friction? UK Teachers' Experiences of Cross-Cultural Teaching in China

Harriet Axbey

School of Education, Durham University, Durham, United Kingdom

harriet.a.axbey@durham.ac.uk

Harriet Axbey is an ESRC funded student in the School of Education, Durham University.

Fusion or Friction? UK Teachers' Experiences of Cross-Cultural Teaching in China

Cultural exchanges between the UK and China have gained in popularity in recent years due to the success of East Asian countries in international tests such as PISA. Chinese 15-year-olds outperform their British counterparts in mathematics, and many practitioners are travelling to China to learn teaching techniques in the hope of raising standards back home. Twenty-six undergraduate trainee-teachers and eleven fully-trained teachers travelled to the South of China to observe mathematics teaching in primary schools. They also took part in the teaching of English, Mathematics, and Science. The aim of the trip was to observe the phenomenon of 'teaching for mastery' as advocated by the National Centre for the Excellence in Teaching Mathematics (NCETM). After the trip, four participants were interviewed via email on their experiences in China. The researcher found that the three trainee teachers and one qualified teacher felt that what was being implemented in the UK as 'teaching for mastery' had little to do with the actual practices in China, where the concept supposedly originated. The participants commented that they had witnessed several teaching methods they would bring back to their own classrooms, but that the UK should be wary of adopting practices without seeing them in the context from which they came. Additionally, participants found many differences in terms of behaviour, class structure and teaching experience, which will be discussed along with the questions: Can we transfer Chinese teaching directly to the UK? And: Is teaching for mastery actually a British construct?

Keywords: mastery, mathematics education, teacher training, China, teaching for mastery

Introduction

The concept of mastery in mathematics was born out of a desire to emulate the success of East Asian countries in traditional assessments (Boylan et al., 2018). Attempts to transfer East Asian mathematics teaching methods to the UK, such as the clustered randomised control trials by Jerrim and Vignoles (2015), appeared to show a modest but positive treatment effect when mastery was implemented. Several teaching exchanges and cross-cultural comparisons such as those by Boylan et al. (2018) and Norton and Zhang (2016) have attempted to examine and explain differences in teaching methods and student attainment. Norton and Zhang (2016) state that Confucian values of mathematics influence the way in which it is taught in China and East Asia, with the subject not needing to be seen as immediately relevant to the learner.

This paper will look at the experiences of trainee teachers and one qualified teacher following on from a trip to China where they observed lessons in mathematics, as well as teaching in a Grade 5 (age 10-11) class themselves.

Twenty-six trainee teachers and eleven qualified teachers travelled to China as part of a fusion programme designed to encourage the sharing of methods between Chinese and UK practitioners. Three of these trainee teachers and one qualified teacher were interviewed via email following the trip, regarding their experiences observing and teaching in Chinese primary schools. Findings suggest that all the teachers witnessed teaching very different from what they were expecting, and two explicitly stated that what they had believed to be teaching for mastery was very different from the reality in Chinese classrooms.

Context

The fusion of Chinese teaching methods into British schools has been a topic of discussion and research for many years now. This interest is based on studies such as those by Jerrim and Vignoles (2015) showing an increase in children's attainment following the implementation of so-called mastery principles. Tyumeneve et al. (2014) also found that mastery had a high positive relation to the ability to transfer mathematics to out of subject contexts. There is also some evidence of Chinese teaching incorporating Western methods; Yang and Li (2017) stated that Chinese kindergartens have been hybridising Eastern and Western curricula to improve provision. However, most of the fusion reported in literature has been on the adoption of Chinese methods into a British context.

The beliefs of teachers can be a major factor in reform of education, and adoption of mastery methods would require the support of teachers, parents and pupils (Boyd & Ash, 2018). Some features of what is called mastery, such as mixed ability settings, are not encouraged by many. Francis et al. (2016) found that parents supported streaming by prior attainment purely because they believed their children would be in the top sets, and that research supporting mixed ability grouping has had no impact on the reality in the classrooms. Yu (2009) examined the differences in the beliefs of Chinese and English teachers, and found that there

were significant distinctions between the way that they viewed teaching and mathematics as a concept. They discovered that teachers had similar beliefs about the importance of using real world examples, but that Chinese teachers placed more emphasis on the importance of the introduction of new concepts and methods, as well as the nature and function of proofs and the role of precise language.

Yu (2009) further found that English teachers tended to have a pragmatics understanding of theory in relation to mathematics, whereas their Chinese colleagues had a more scientific understanding. Boyd and Ash (2018) also explored teacher beliefs, following the implementation of a textbook-based 'Asian mastery' approach. They found that teachers' cultural beliefs changed, and that they began to believe in malleable mathematics intelligence as opposed to a fixed intelligence, changing their commitment to grouping pupils by prior attainment (Boyd & Ash, 2018).

Norton and Zhang (2016) did a cross-cultural study of Australian and Chinese teachers, they made several salient discoveries relevant to the findings of this study. They stated that Chinese teachers held a privileged position in society, and that this allows teacher training in China to attract the most knowledgeable candidates. They commented that many Australian teachers struggled with the kind of mathematical content that they would be expected to teach, whereas Chinese teachers showed mastery. However, they also commented that the relevance of the difference in students' scores could be dismissed due to the fact that Chinese students spent more time studying, and this would therefore have an impact on their mathematics attainment (Norton & Zhang, 2016). They further dismissed the children's scores as being due to the international tests being comprised mostly of low-order questions, which they claim Chinese students excel at (ibid.).

Miao et al. (2015) compared teaching in China and England and found that Chinese teachers scored higher on effective teaching measures, as well as pupils outscoring their English peers. However, the head teachers in the chosen schools were able to choose which teachers were studied (Miao et al., 2015). These sort of comparisons, and those by Norton and Zhang (2016), between Chinese, and English and Australian teachers, attempt to describe the differences between teachers, without taking into account the similarities. They also often negate to take into account the number of other factors that could impact upon attainment scores, and whether indeed these attainment scores are a relevant measure of teaching success, or pupil ability.

The teaching profession is seen very differently in China and in the UK, with teachers holding a privileged position in society (Norton & Zhang, 2016). Furthermore, teachers in China are experts in their fields, many holding a Master's degree in their particular field, and are expected to engage in intensive ongoing school-based professional development (Boylan et al., 2018). Teachers in Chinese primary schools were observed to teach around two or three forty-minute lessons a day each, repeating the same lesson to each class in a grade.

Conversely, in the UK, primary teachers are facing ongoing pressure in many areas, and are expected to teach all subjects, all day, with only a few hours a week dedicated to planning and development. Primary teachers therefore cannot be specialists in every subject that they teach, and hold a very different place in society than their Chinese colleagues. (Boyd & Ash, 2018).

Methods

Originally the researcher planned to conduct face-to-face interviews, however, due to the undergraduate trainee teachers leaving the area for their Easter holidays, email interviews had to be undertaken instead. As an asynchronous mode of interviewing, email has many benefits as well as limitations (Bryman, 2016). Answers tend to be more thorough and grammatically correct, as the respondents have more time to consider their answers; additionally, email is a useful and unique tool when nonverbal and paralinguistic cues are not needed for evaluation (Curasi, 2001; Ratislavová & Ratislav, 2014). For this study, what respondents said was considered more important than the need to look at their tone and movements.

Burns (2010) stated that the email interview method enriches the array of investigatory tools available for qualitative social researchers today, although they will never be a replacement for face-to-face interviews (Ratislavová & Ratislav, 2014). For this study, participants were recruited from a group of thirty-seven people who took part in a university-led trip to China to visit, and teach in, primary schools. Out of this group, twenty-six were undergraduate trainee teachers and eleven were qualified teachers. All members of the group were contacted with a request to provide their university email if they wanted to take part in the study. The researcher received seven responses, and three students and one qualified teacher responded to the subsequent email with their answers to the questions.

Participants were emailed a consent form and information sheet, with five questions and a space for adding additional comments. The questions were designed to be as open-ended as possible, and asked participants about their experiences of the Chinese education system, their views on the teaching they had witnessed, and the methods or strategies they would take back to use in their own classrooms in the UK. Although the response rate was relatively low, this is expected of the email interview method (Bryman, 2016).

As Bryman (2016) stated that participants often lose momentum when they are expected to reply to more than one email, only one communication was expected of participants. The email method proved effective as an interview method, as social interaction was not required for the research, and standardisation of interview was not needed (Ratislavová & Ratislav, 2014). This method is further effective when anonymity is beneficial (Ratislavová & Ratislav, 2014), as it was in this case, as trainees preferred their comments to be kept anonymous. The method also allowed for participants to interpret the questions in their own way, and to go off on a tangent if they wanted to (Bryman, 2016). This helped with the

research as it allowed the researcher to see what was important to the participants, although email interview often results in less detail than a face-to-face method (Bryman, 2016).

Findings

The overall finding was that the trainees and qualified teacher believed the UK should be wary of transferring the methods used in Chinese schools to schools in the UK, and that teaching for mastery is not necessarily the practice used in China. Additionally, trainees commented on the atmosphere and ethos of the schools they visited, mentioning many positive attitudes towards both learning and teaching. This echoes Norton and Zhang (2016) who stated that Chinese teachers held a privileged position in society.

Ethos and atmosphere

The participants witnessed several teaching practices that they would transfer back to their own classrooms, as well as some practices they would adapt, or not use at all. The qualified teacher commented that they could see that the UK has a lot to learn from international practice, but that the UK pedagogy also has a lot to commend it. They noted that the last point was 'too often overlooked and we are in danger of losing the good practice we have developed by simply replacing it with practice from another country.'

All participants commented on the positive attitudes towards school and learning held by teachers and students in the schools they visited. One participant, when asked what strategies they might use from observed lessons, replied that it was hard to answer the questions because 'a lot of the good teaching methods and strategies come about because of the money allocated to teacher training and to schools in China, alongside the perceived importance of education.' Another participant, in response to a similar question, stated that they would encourage a classroom culture where talk is encouraged, expected and valued as a learning tool by all.

Teacher subject knowledge was also mentioned by participants; in China, teachers teach one subject e.g. English, or Chinese, or Mathematics, and so are experts in their particular field. Where we do have this system in most secondary schools in the UK, it is not normal practice in primary education. The qualified teacher stated that they had witnessed 'unbelievable' music and art lessons, which could not be replicated in the UK due to lack of teacher subject knowledge and accessibility of resources.

Representations

Concrete, pictorial and abstract manipulations are a key feature of teaching for mastery as defined and demonstrated by Jerrim and Vignoles (2015). Observation of the use of representations and manipulatives was mentioned by all participants, confirming the belief that these form a large part of the teaching in Chinese schools. One participant observed that children used Lego as a manipulative to aid calculations within word problems that involved

multi-step addition and subtraction a part of a narrative. They commented that this use of narrative and concrete representations led the children into the learning rather than presenting them with the calculations up front.

Another participant mentioned the use of concrete manipulatives such as plastic sticks, straws and elastic bands, which all children are provided with for their own use in individual small boxes. This allowed children to use these manipulatives at home and in lessons, to support their learning and understanding of mathematical concepts. The trainee commented, however, that the UK is investing a lot of money into branded concrete manipulation resources (such as dienes and numicon), based on the belief that mastery uses these. However, they stated that the resources in Chinese schools such as these small boxes and the teacher using a tissue box to explain shape, looked much cheaper and simpler, but led to the same outcome, suggesting that the UK is investing unnecessary money in mastery resources that are not actually featured in the successful East Asian countries.

On the subject of commercial schemes such as those which produce concrete resources, one participant stated that the practice of teaching from schemes such as White Rose, Power Maths and Maths No Problem can 'stifle teacher and pupil creativity…facilitate gaps in pupil knowledge and deskill teaching staff.' This goes against recommendations by the NCETM (2016) to use such textbook-based approaches.

Writing

Regarding formal written work, one trainee stated that they would evaluate how much of this kind of work students in their class were doing, as they observed the students in the Chinese classrooms doing very little of this sort of work. Children were observed working practically and collaboratively with very little writing, and sometimes none at all. They believed that this was beneficial in keeping the children engaged and on task with their work. In the UK, children's written work is constantly evaluated as a product of a lesson through book scrutinies and leadership walkrounds. The quality and quantity of children's work is seen to be a reflection of the quality of teaching the child has received. This is not the case in the schools visited in China; all children worked out of workbooks and textbooks, and therefore, aside from mistakes and differences in handwriting, every child, every year, will complete and same work in the same order.

However, the practice of completing most written work at home is not beneficial in all ways; one trainee commented that they would prefer to mark their students' work during the lesson where possible, to allow children to re-evaluate their answers and learn from their errors immediately. They observed very little assessment for learning during their time in China. One trainee also questioned how the use of textbooks allowed for adapting subsequent lessons to the needs of the pupils. Although they believed the textbooks gave a strong structure to the curriculum delivery, they did not agree with moving all children along at the pace as determined by the textbook. They stated that they believed the UK strategy of

adapting lessons for the needs of the pupils benefits the less able children in their learning much more than the Chinese approach they witnessed.

Differentiation and whole-class teaching

With regards to the concept of mastery, participants stated that the British idea of mastery was very different from the Chinese reality. One trainee recalled a conversation with their headteacher, who stated that they did not expect all children to progress at the same pace, but that the mathematically talented children should be taught different material in order to access higher levels of education in mathematics. This directly contradicts the principles of mastery as it is implemented in the UK by the NCETM (2016) and described by researchers such as Jerrim and Vignoles (2015).

Furthermore, one trainee commented that Chinese teaching seemed to have an over-reliance on whole-class teaching, which led to only a small minority of pupils volunteering information and answers, when other pupils took a more passive role. Common practice in the UK would involve a great deal of paired and table-based work, leading to a more collaborative approach. The trainees witnessed a very different approach in China, which they believed was detrimental to the students' learning. Whole class teaching is generally encouraged in East Asia, and therefore is seen as a part of teaching for mastery (NCETM, 2016). Furthermore, Miao et al. (2015) found that pupils performed better in classrooms where teachers allocated more time to interactions with the whole class, and less time with individual pupils or groups. However, they also made it clear that they did not believe this should mean there should be no group work in the classroom (Miao et al., 2015).

Boyd and Ash (2018) believe that the Western method of teacher demonstration followed by practice is a result of cultural myths such as the belief in natural talent. However, the findings in this paper show that belief in mathematical talent is not necessarily a Western myth, as it is also a belief among some teachers in China, leading them to deliver different content to those pupils. Boyd and Ash (2018) also stated that teachers who followed the mastery curriculum began planning their lessons towards subject knowledge. However, the trainee teachers observed very little planning, with most of the lesson pre-planned based on the structure of the textbooks.

Variety in calculation

One participant stated that, following their observations of Chinese mathematics teaching, they would advise newly qualified teachers (NQTs) to use a variety of calculation strategies, as this would give children a chance to find and use methods best suited to their style of learning. This echoes Yu (2009) who believed that the Chinese method of teaching mathematics involved a lot of logical reasoning, and application to other areas, and a variety of calculation strategies would help with that. One trainee witnessed an effective lesson where the teacher used partitioning alongside long multiplication, they stated that this variation helped children to consider their own metacognitive processes, and allowed them to

decide on the best approach for a particular problem. This is an example of where using a variety of calculation strategies was used effectively to support children's learning, and participants stated they would use such an approach in their own classrooms.

Behaviour

Participants commented that they were surprised at the amount of low-level disruption in the classes. One participant stated that they expected a much more disciplined classroom, but experience chatter and disruptive behaviour such as looking in bags and drawing while the teacher was talking. Furthermore, one trainee witnessed more serious behaviour problems, where two children sustained broken bones; the teacher in this case blamed their behaviour on the lack of time they had to play outside. The trainee believed that this lack of time to play had a detrimental effect on the children's overall education.

Although some behaviour management techniques were witnessed, for example one teacher played music to get the class's attention, overall, trainees noticed a profound difference in behaviour from what they expected, and what they were used to in the UK. Researchers such as Boylan et al. (2018) talk of exchanges, but mostly just discuss what the UK can learn from China. However, as one participant stated, there are some areas of UK education, such as behaviour management, that are more successful than the Chinese methods, and therefore an exchange should involve a two-way trade of information and best practice.

Conclusion

This paper does not attempt to state whether one country delivers mathematics better than another, but it does set out the major differences between what is believed to be teaching for mastery in the UK, compared to what is common practice in China. Many attempts have been made over the years to improve the teaching of mathematics, a considerable amount of change occurring after the publication of the Cockcroft Report (1982) which put a large emphasis on verbal exploration and reasoning (Boyd & Ash, 2018).

Mastery in mathematics is a concept of teaching that was brought over from East Asia and implemented by practitioners such as the NCETM (2016). Researchers such as Jerrim and Vignoles (2015) and Miao et al. (2015) have explored how these techniques can be transferred to the UK. However, fusion of two curricula must involve a two-way exchange, and any adaptation of another culture's practices must be done with caution. Furthermore, Norton and Zhang (2016) explained that China could look to the West for teaching reasoning, but that it should not forget that it is good at the fundamental basics of mathematics.

On the 2019 trip, the trainee teachers gained a great deal from their experiences in Chinese primary schools, they developed their teaching techniques in many areas of the curriculum, made connections with their Chinese colleagues, and examined the extent of whether what they had been told was teaching for mastery was actual practice in China. However, they also

saw many differences in practices which they believed were not indicative of a mastery curriculum as they had seen it in the UK. This echoes Boylan et al. (2018) who explained that the mastery innovation in England was not based on a systematic review of effective practice in mathematics, but solely from a desire to emulate the success of Shanghai and Singapore. Additionally, these findings suggest that China could learn from the UK in regards to aspects of behaviour management, as Chinese teachers are extremely informed in subject knowledge, but this must be supported by effective classroom management strategies.

Many practices, such as the use of concrete, pictorial and abstract representations, and varying the calculation strategies shown to children, the trainee teachers would happily adopt in their classrooms. This supports the view of Boylan et al. (2018) who believed that many of the specific practice of mastery, if considered individually, had the potential to improve attainment. However, as one participant stated, there are also many aspects of Chinese teaching that would not work in the UK, and the UK should be wary of imparting an entire style of teaching without considering the cultural differences involved. It is possible, if teachers in the UK held the privileged position in society that Chinese teachers do, and if UK schools had the funding to provide the resources that some Chinese children have access to (textbooks, extra-curricular activities, large facilities), many aspects f UK teaching would be very different.

One participant stated that they believed more research was urgently needed to re-evaluate the currently-used interpretation of teaching for mastery, and to determine the specific features of the teaching method which have a positive effect on pupil attainment. The researcher would agree with this, and additionally echo the call from Francis et al. (2016) who believed that mixed ability grouping needed a basis in scientific truth based on randomised control trials. More research is therefore encouraged into the differences between teaching for mastery in the UK, and practice in China, and an effort to produce further data supporting the teaching of children in mixed-ability settings.

References

Boyd, P. and Ash, A. (2018). Mastery mathematics: Changing teacher beliefs around in-class grouping and mindset. *Teaching and Teacher Education*, 75, 214-223.

Boylan, M., Maxwell, B., Wolstenholme, C., Jay, T. and Demack, S. (2018). The Mathematics Teacher Exchange and 'Mastery' in *England: The Evidence for the Efficacy of Component Practices. Education Sciences*, 8(4), 202.

Burns, E. (2010). Developing email interview practices in qualitative research. *Sociological Research Online*, 15(4), 1-12

Cockcroft, W. (1982). *Mathematics counts. Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr WH Cockcroft*. London: Her Majesty's Stationery Office.

Curasi, C. (2001). A critical exploration of face-to-face interviewing vs. computer-mediated interviewing. *International Journal of Market Research*, 43(4), 1-13.

Francis, B., Archer, L., Hodgen, J., Pepper, D., Taylor, B. and Travers, M. (2016). Exploring the relative lack of impact of research on 'ability grouping' in England: a discourse analytic account. *Cambridge Journal of Education*, 47(1), 1-17.

Jerrim, J. and Vignoles, A. (2015). The link between East Asian 'mastery' teaching methods and English children's mathematics skills. *Economics of Education Review*, 50, 29-44.

Miao, Z., Reynolds, D., Harris, A. and Jones, M. (2015). Comparing performance: a cross-national investigation into the teaching of mathematics in primary classrooms in England and China. *Asia Pacific Journal of Education*, 35(3), 392-403.

National Centre for Excellence in the Teaching of Mathematics (2016). *The Essence of Maths Teaching for Mastery*. [online] National Centre for Excellence in the Teaching of Mathematics. Available at: https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf [Accessed 14 Apr. 2019].

Norton, S. and Zhang, Q. (2016). Primary mathematics teacher education in Australia and China: What might we learn from each other?. *Journal of Mathematics Teacher Education*, 21(3), 263-285.

Ratislavová, K. and Ratislav, J. (2014). Asynchronous email interview as a qualitative research method in the humanities. *Human Affairs*, 24(4).

Tyumeneva, Y., Valdman, A. and Carnoy, M. (2014). How Well Do You Need to Know It to Use It?. *SSRN Electronic Journal*.

Wang, W. (2015). Teaching English as an international language in China: Investigating university teachers' and students' attitudes towards China English. *System*, 53, 60-72.

Yang, W. and Li, H. (2017). A school-based fusion of East and West: a case study of modern curriculum innovations in a Chinese kindergarten. *Journal of Curriculum Studies*, 50(1), 17-37.

Yu, H. (2009). A comparison of mathematics teachers' beliefs between England and China. *Research in Mathematics Education*, 11(1), 83-84.

Corpus-based analysis of lexical cohesion in Chinese postgraduates' English academic writing and its pedagogical implications

Xuanhong Guo

School of Education, Durham University, Durham, United Kingdom

xuanhong.guo@hotmail.com

Xuanhong Guo is a China Scholarship Council funded PhD student in the School of Education at Durham University.

CSC Funded PhD Project

Corpus-based analysis of lexical cohesion in Chinese postgraduates' English academic writing and its pedagogical implications

As a key feature in the creation of coherent texts (Tanskanen, 2006), lexical cohesion is of importance for students' academic performance. Chinese students have been identified as lacking awareness of lexical cohesiveness in English academic writing (Zhang, 2018). In order to inform pedagogy in English for academic purposes (EAP) for these students, this paper used a corpus-based approach to conduct qualitative analysis of lexical cohesive devices used in Chinese postgraduates' writing at a UK university. A framework for the analysis of lexical cohesion was developed in two corpora, incorporating a new subcategory of lexical cohesive device alongside modifications of existing categories. Analysis of the corpora identified homogeneities of lexical cohesion such as context sensitivity, dominant use of repetition, and use of modifiers to indicate lexical cohesive relations, suggesting the value of context-based pedagogy and the need to teach lexical cohesive devices with appropriate exemplars.

Keywords: lexical cohesion; Chinese EFL learners; EAP pedagogical implication

Introduction

Cohesive devices help create the connectedness of texts (Halliday & Hasan, 1976, p. 2), which contribute to developing the meaningfulness of texts and impact upon communicative effectiveness (Tanskanen, 2006, p. 1). Based on the forms of expressing cohesive relations, cohesion is divided into grammatical cohesion and lexical cohesion (Halliday & Hasan, 1976, pp. 5-6). As with this mainstream practice of putting models of cohesion to use, this paper explores the features of inter-clausal lexical cohesive devices used in creating cohesive academic written texts within a specific academic discipline, in order to provide English for Academic Purposes (EAP) pedagogical implications to Chinese students as English as a Foreign Language (EFL) learners who have been identified in general as lacking awareness of lexical cohesiveness in English academic writing (Zhang, 2018).

Lexical cohesion is defined as relations in which two or more lexical items connect to each other and to other cohesive devices, in order to "build the continuity of the text" (Flowerdew & Mahlberg, 2009, p. 1), in general consisting of two categories:
reiteration and collocation. Other sub-categories within these two categories vary among models of lexical cohesion in previous studies. Therefore, this study also developed its own classification of lexical cohesion for this specific research context.

In terms of the application of cohesion in teaching and learning, researchers have studied the use of cohesive devices across a variety of EAP contexts as well as in both native and non-native writing in foreign and second language settings (e.g. Ong, 2011; Sinicrope, 2007; Zhang, 2000). These researchers without exception believe that lexical cohesion is vital in textual cohesion, playing a crucial role in text interpretation whether the language user is a native or a non-native speaker (henceforth NNS) of English. However, surprisingly, few

studies solely focus on lexical cohesion in higher education students' actual on course writing when such studies have the potential to inform EAP pedagogy. Furthermore, there have not been any such studies in the UK higher context, though Chinese students are already the largest subgroup of international students in the UK academic settings (British Council, 2017). In order to remedy these limitations, this study focuses on a detailed analysis of lexical cohesion in two corpora of excerpts from Chinese students' module assignments and MA dissertations submitted to MA TESOL and MA Applied Linguistics for TESOL programmes at a UK university.

Methodology

52 module assignment samples (17,538 words) and 45 dissertation excerpts (19,148 words) from 9 dissertation texts were collected. Then, due to context sensitivity of lexical cohesion (Xi, 2010, p. 143), a suitable model of lexical cohesion was developed based on previous studies for the manual analysis of lexical cohesive devices used in the corpora (see Table 1). Furthermore, qualitative analyses of specific lexical cohesive relations identified in the corpora were conducted. The reason for adopting the manual analysis method was that lexical cohesive relations are based on semantic grounds, which cannot (at least at the moment) be identified with the assistance of such concordance tools (Hoffmann, 2012, p. 101). The analytical table is demonstrated in Table 2.

Table 1. Classification of lexical cohesion in the present study

Category	Definition
Repetition	This category has been divided into two types in the present study: simple repetitions (i.e. the same lexical item) and complex repetitions with derivational variations or grammatical changes.
Synonymy	The meanings of the lexical items can be interpreted straightforwardly or contextually as being similar.
Antonymy	The meanings of the lexical items can be interpreted straightforwardly or contextually as being opposite.
Hyperononymy /hyponymy /meronymy	A hyperonymic relation is defined as the "relation which holds between a more general, or superordinate, lexeme and a more specific, or subordinate, lexeme" (Hoffmann, 2012, p. 90). There are two types of hyperonymic relations in the present study: kindwhole and part-whole relations. Hyponymy and meronymy refer to whole-kind and whole-part relations respectively.
Signalling nouns (SNs)	The category of SNs in lexical cohesion includes SN-like nouns, especially the general nouns (Halliday & Hasan, 1976), the shell nouns (Schmid, 2000) and the discourse-based signalling nouns (Flowerdew & Forest, 2015). The general features of SNs are: SNs have 'semantic generality', can be used as 'containers' for more specific contents, and can encapsulate complex information into highly condensed concepts with simple lexical items.
Identity	The full name for this category is 'other relations with identity of reference' (identity), dealing with other co-referential relations which are not included in the existing categories.
Collocation	Collocation comprises the semantic relations between lexical items beyond the clause which are generally known to be part of some larger objects or situations and frequently co-occur.

Table 2. An analytical table for the manual analysis of lexical cohesion (from sample F9 in the failed group)

text	repetition	synonymy	hyperonymy	hyponymy	meronymy	SNs	identity	antonymy	collocation
Part One Target Group									
This **assessment** tool aims at a **group** of **business-major students**	group		business – major				target group – students		
who are joining an **English** **training** program,									students – training program
after which they are going to become **interpreters** and **bilingual** **sales agents** in an **international furniture fair**.							Business-major students – interpreters and bilingual sales agents		English – interpreters business – sales agents

As it can be seen in Table 2, the left-hand column of the analysis sheet specifies the clauses in the sample while the other columns demonstrate the lexical cohesive categories of the relations and their corresponding cohesive pairs found within the text. The analysed lexical items in column one (left) were coloured or highlighted in the analysis for the convenience of retrieval and analysis (see Table 3 for the highlighting system of the manual analysis). Each pair recognised was allocated to its corresponding category column, and a hyphen was introduced between elements in that pair. The cohesive pairs in a clause were placed in the cells which were in the same row of that clause cell. For example, in table 2 above, because *interpreters* and *sales agents* were in the same clause, the collocation pairs *English – interpreters* and *business – sales agents* were placed in the same cell in the same column.

Table 3. Highlighting system in the present manual lexical cohesion analysis

lexical cohesion category	example
bold is for repetition	**speaking – speaking**
orange is for synonymy	test – assessment
dark red is for hyperonymy	English – language
green is for hyponymy	language – English
red is for meronymy	English skills – listening
highlight grey is for signalling nouns ("//" is used to divide clauses)	Literature on how to respond to the inappropriateness of arguments clearly in teacher feedback is scant, //and pedagogical ESL suggestions for writing teachers on how to deal with such issues could be a subject for further studies.
blue is for identity	test – element
purple is for antonymy	learning – teaching
highlight yellow is for collocation	lecture – student

Results and discussion

The appropriate examples (see Table 4 below) for each category will be displayed respectively at first to demonstrate the proper use of each lexical cohesive device in texts. EAP pedagogical implications will be provided following the examples to give further suggestions for Chinese students regarding their study of using lexical cohesion in academic writing.

Table 4. Examples of lexical cohesive categories in the present study

(1)	there has been a lot scholars and **researchers**// who **researched** …// Most of these **researche[r]s** …// Thus, **research** questions … (from sample D6C3M) (Grammatical typos in the original sample text are corrected in "[]" where applicable)
(2)	the learners// The course will be taken by 10 to 12 Chinese young adults (late teens or early twenties) as part of their preparation for meeting the level of English language proficiency required for admission to UK universities.// All of the students wish to achieve band score at least 5.5 in IELTS speaking … (M12P)
(3)	The second part reflects the approach of task-based analysis, which helps to ensure the course to possess a high degree of real-life relevance … // … they can answer the questions according to the occasions where they cooperate with Chinese clients. This could also be conducive to figure out the tasks they are likely to carry out with foreigners. (D6)
(4)	As an ending of the course, the department of human resource (HR) in the company wants to find out the participants' learning outcome.// Therefore, an assessment tool is required to design not only for evaluating achievement relevant to the short-term program, but also for deciding the final list of employees. (P5)
(5)	The students may even find it difficult to notice the errors when speaking English. Thus the perceived needs of students are, […] explicit teaching of the thinking difference which is accountable for the errors, a raised awareness of the thinking processes and the thinking habits in speaking English to encode time information in verbs and gender information in the third person pronouns.

Therefore, the syllabus is designed to achieve the goal of the "entrenchment" of a chain of thinking processes that can generate correctly-tensed and gender-referred speaking English. (M1P)

(6) There are totally 22 students in class,// among which 16 are girls// and 6 are boys. (D9P)

(7) But even where performance test materials appear to be very realistic compared to traditional paper-and-pencil tests, it is clear that the test performance does not exist for its own sake. However, it is necessary to have a procedure that is fair to all candidates, and elicits a scorable performance, even if this means involving the candidates in somewhat artificial behaviour. (D2)

(8) It represents a particular realization of communicative language teaching.// … they can enjoy the activities and create more active learning atmosphere with satisfactory effect. (M7P)

(9) However, the uneven representation of values could be a feature of the data set and the individual speakers, … // … This unpredictability in terms of difficulty is exacerbated by the individual listener who has their own strengths and weaknesses. (D8C5C)

(10) This dissertation is a study focusing on the exploration of law students' past English learning experiences and spoken English issues in seminars through their reflection on seminar learning in the LLM programme. … Ethnographic methods offer this study a holistic approach to … systematically document[ing] the influences of students' background on their seminar learning in a rich, contextualised detail with the aim of suggesting proper measures to deal with language issues in law seminars. (D7C1I)

(11) Chapter One Introduction//… With regards to peer feedback in this dissertation, it refers to the activity where students read each other's essay and then express not only negative criticism but also supportive and appreciated evaluation. (D13C1)

(12) because essays and reports are usually regarded as the most popular assessment forms in academic courses in western countries.// They may also have exams but compared with short exam answers,// it will cost more time like several weeks to write essays as assignments for course work. (D11C1I)

(13) Supplementary education, also known as "shadow education" or "private tutoring" has been expanded rapidly in the globe since this century. This phenomenon has first been developed in East Asia and has become externally visible throughout Asia as well as in other world regions in the present days. (D12C1I)

(14) We can see this through an exquisite job done by Quirk et al. , which categorized number classes of nouns mainly into four groups// … Nouns in group (A) are occurring only in singular form, which include (Aa) mass nouns such as gold, music, (Ab) abstract adjective heads like the unreal, and (Ac) some proper nouns like Henry, the Thames.// On the contrary, nouns in group (B) are occurring only in plural forms, which can be distinguished into five subgroups… (D2C2L)

(15) From the information which is attained from the interviews, the overview of participants['] perspective on peer feedback is that// five of the eight students (Students A, B, C, D, E and Student G) felt that this activity was helpful. (D13C4F)

(16) With its development, the popular and dominant ELT methods in China are grammar translation and audio lingual,// [and] they are 2 ways that make great contribution to language teaching. (D6C1I)

(17) The law students […] have at least two seminars in a week […]// The size of seminars varies:// some seminars are quite small, only consisting of 6-8 people … (excerpt 4)

(18) For example, if only grade one is analysed, perhaps the higher grades start to introduce implicature// because designers suppose these freshmen do not acquire sufficient pragmalinguistic knowledge … (D14C3M)

(19) This means that more university applicants will choose IELTS examination to attain a place in the course.// If students want to apply for a Tier4 general student visa, their IELTS overall score has to reach the band 5.5 and above… (D4C1I)

(20) This assessment tool takes the form of a speaking test, devised as a progress test based on the syllabus// which aims to entrench the automatic cognitive processing to encode time information into verbs and gender information into third personal pronouns for Chinese learners of English.// Moreover, … the task also purposes [proposes] an evaluation of the extent// to which the students can use the target language to communicate and co-construct conversation regarding familiar topics. (D1)

(21) it is common that researchers prefer questionnaires rather than interviews,// since perhaps the former tools can be used to attain information from a large number of participants … (D13C3M)

(22) Chapter Two Literature Review// This chapter will look at relevant research based on theoretical research as well as the investigation related to peer feedback from students' perspective, including Asian students, Chinese and Japanese and European learners from Spain. (D13C2L)

Repetition

The findings suggest that Chinese students tend to use simple repetitions rather than complex repetitions, indicating the necessity of introducing complex repetitions to students, in order to raise their awareness of avoiding overusing simple repetitions which may make texts seem uninteresting.

In example 1, research-stemmed repetitions form a chain with three repetitive pairs: *researchers – researched – researchers – research*. Sharing the same word stem – *research*, these pairs may potentially be interpreted as being lexically cohesive by readers. On the other hand, these research-stemmed lexical items are slightly different in forms, which helps to reduce the monotony of the text, representing a useful example to teach the use of complex repetitions.

Synonymy and antonymy

The categories of synonymy and antonymy are combined for discussion because of similar divisions based on two criteria: whether the lexical items in a synonymous or antonymic pair belong to the same word class; or whether the meanings of the lexical items can be interpreted straightforwardly as being similar or opposite. Specifically, synonymy/antonymy is divided into synonymy/antonymy in the traditional sense and near-synonymy/antonymy.

Example 2 demonstrates a synonymous relation in the traditional sense between *learners* and *students*. *Learners* is replaced by *students* in the following clause. It is not difficult to interpret their synonymous relation as they have similar word meanings: *learners* represent people who are learning something (OED Online, 2019), and *students* denote people who are learning at school (OED Online, 2019). Furthermore, they share the same referent, *10 to 12 Chinese young adults*. These two points make it clear that *learners* and *students* form a synonymous relation in example
2.

By contrast, examples 3 and 4 demonstrate near-synonymous relations from two perspectives. In example 3, although both items express the meaning of providing something good to make other things happen (OED Online, 2019), *helps* (noun) and *conducive* (adjective) belong to different word classes. According to criterion one, they are regarded as near-synonyms in example 3. In example 4, *outcome* and *achievement* form a near-synonymous relation because of sharing the same referent, the participants' learning results; and their contextually similar meanings: *outcomes* refers to the participants' learning result of the course, and *achievement* refers to the good result of the participants' learning in the course, both denoting the participants' learning results. However, *achievement* adds a positive feature to the results while *outcomes* expresses a neutral meaning. Therefore, they are regarded as near-synonyms.

In terms of antonymy, in addition to the introduction of the two types of antonymy mentioned above, this category can also be categorised into four subcategories according to the opposite

relations between two lexical items in a antonymic pair, three of which being recognised in the present study and discussed with examples below.

The first sub-category is complementary antonymy, denoting non-gradable binary contrast between two antonyms.

In example 5, *errors* and *correctly* are a near-antonymic pair as *errors* is a noun while *correctly* is an adjective, and they express contradictory meanings in this context: *errors* refers to the inappropriate contents in students' spoken English, whereas *correctly* in *correctly-tensed* denotes the appropriate use of verb tense in students' speaking English, which in general also refers to the contents in students' spoken English. Therefore, *errors* and *correctly* are contradictory in meanings here, demonstrating the feature of antonymy as lexical cohesive devices that two lexical items from different word classes can form near-antonymic relations when expressing contradictory meanings in a specific context.

The pair *girls – boys* in example 6 is much more straightforward than *errors – correctly* regarding the contradictory meanings of the lexical items. The former items are semantically regarded as contradictory in general, while the contradictory meanings of the latter need interpretations in specific contexts.

The second sub-category is contrary antonymy, referring to relations between gradable antonyms. In example 7, *realistic* and *artificial* constitute a contrary pair, as the former is related to the reality or authenticity, while the latter means not authentic. The interesting point is the use of the surrounding indicators implying the comparable sense between *realistic* and *artificial*, such as *very, compared to* and *somewhat*, which make it clearer that *realistic* and *artificial* are gradable in terms of their contrary meanings, providing a good suggestion for teaching gradable antonyms with such indicators in the co-texts.

The third sub-category is converse antonymy, which typically exists in two situations: procedural verbs and nouns expressing reciprocal social roles. In the first situation, the actions expressed by the verbs are involved in a unidimensional movement from two perspectives: that of the source and that of the goal (Murphy, 2003). In the second situation, one social role cannot exist without the other. In example 8, *teaching* and *learning* form a near-antonymic converse pair. Firstly, the two items are near antonyms because *teaching* is a noun while *learning* is an adjective. Further, the items express the same procedure from teachers' perspective (the source) and students' perspective (the goal) respectively. Therefore, *teaching* and *learning* are converse antonyms in this context. In example 9, *speakers* and *listener* are two social roles which are interdependent. As Crystal (2008) commented, there is symmetry of dependence in the reciprocal social role. The object of the *speakers* is the *listener*, and the object of the *listener* is the contents that the *speakers* say. Both *speakers* and *listener* cannot exist without each other. The two examples show the interdependence between two lexical items in a converse relation in two typical situations,

which gives practical suggestions to design appropriate examples for teaching converse antonyms.

In general, teachers can illustrate each type of antonyms with exemplars such as the examples above at first, and design matching activities where learners need to match listed antonymic relationships with individual sentences in which different antonyms are used.

Hyperonymy, hyponymy and meronymy

Hyperonymy, hyponymy and meronymy are included under the umbrella term 'superordinate relations'. Hyperonymic relation is divided into two types. Example 10 refers to a kind-whole relation, *English – language*, while example 11 denotes a part-whole relation, *Chapter One Introduction – this dissertation*.

The relation between hyponymy and hyperonymy is that a hyperonym consists of several types of hyponyms. In example 12, *assessment forms* and *exams* form a hyponymic relation as the more general item *assessment forms* appear before the more specific item *exams*. A co-hyponymic pair is also identified within the same category:

exams and *essays* are two kinds of *assessment forms*, and therefore are co-hyponymic.

The relation between meronymy and hyperonymy is that a hyperonym is made up of several parts regarded as meronyms. In example 13, *the globe* is the whole entity while *East Asia*, *Asia* and *other world regions* are parts of *the globe*. Therefore, the meronymic pair is *the globe – East Asia/Asia/other world regions*. In example 14, *group (A)* and *group (B)* are parts of the four groups of 'number classes of nouns' in sample D2C2L. As these two items occur in two clauses, *group (A)* and *group (B)* are regarded as co-meronyms. It is noticeable that the hyperonym of *group (A)* and *group (B)* also appear in the surrounding clause, which is *the groups of number classes of nouns*. This indicates that the hyperonym and its meronyms co-occur in certain contexts, which gives the EAP implication for teaching meronyms that providing the hyperonym of the co-meronyms can contribute to readers' correct interpretation of the comeronymic relation between two lexical items in texts.

It is recommendable for teachers to introduce these superordinate cohesive devices in EAP classes with examples as the six ones above as they are not often used by Chinese students (Zhang, 2000), probably with fun activities, such as crosswords of a set of hyperonyms with their corresponding hyponyms and meronyms.

Signalling nouns

The analysis of SNs includes its surrounding grammatical structures or modifiers because the surrounding elements contribute to confining the contextual meaning and signifying the co-

referential function of signalling nouns, and therefore guiding the readers to decode the complex information that the SNs encapsulate.

In example 15, the first SN is *activity* in the structure of '*this* + SN' which refers back to *peer feedback*. The use of the determiner *this* indicates that *activity* is used as an anaphor to replace the previous more specific nominal phrase *peer feedback*. The second SN is *overview* in the structure of '*the* + noun + *of*'. The referent of *overview* is the whole succeeding clause *five ... helpful*. The usage of the SN *overview* is a demonstration par excellence of the encapsulating function of SNs by summarising a complicated stretch of text into a smaller nominal phrase, which shows a more sophisticated way of expressing the lexical cohesiveness between segments in the same text than the use of repetitive devices.

As for example 16, what is worth mentioning is the use of a pre-modifier, cardinal number *2*, prior to the SN *ways*, which has another function of describing a specific characteristic of *ways*.

In general, when used as anaphorical cohesive devices, SNs are normally preceded by determiners, such as *these*, or other modifiers, such as cardinal numbers. Another feature regarding using SNs as lexical cohesive devices is that the contextual meaning of SNs is dependent on their referents, while the SNs express more general meanings of the referents. This feature is demonstrated explicitly when GNs are used in texts, displayed in example 17. *People* is a GN which refers back to *the law students*. The use of *people* as a GN provides a valuable pedagogical implication regarding the teaching of GNs. In example 17, the contextual meaning of *people* is provided by the prior use of *the law students*, which suggests that GNs should be used when it is clear for the readers to decode their contextual meaning. Otherwise, the overuse of GNs may result in the vagueness of the text and disconnections between segments in texts, which may further create difficulties for readers' comprehension of the overall text (Wu, 2010). Overall, the use of SNs in texts is context-based. Related teaching activities device need to involve exercises with examples.

Identity

In Halliday and Matthiessen's study (2014), some lexical items which form cohesive relations in the repetition and synonymy categories have been described as having "identity of reference" (p. 645) which is the source of the name for this new category here. 'Identity' is used in a double sense: 'identity' denotes that the lexical items in one pair share the same referent; and the lexical items can demonstrate different identities of the same referent.

Example 18 includes an identity pair *grade one – these freshmen*, in which *grade one* refers to the whole group of students who attend the first level of classes at school. This interpretation of *grade one* is supported by the anaphorical use of *these freshmen* in the succeeding clause, as the meaning of *freshmen* is first-year students (OED Online, 2019), and

the determiner *these* indicates that *freshmen* is used as an anaphor to refer back to *grade one*. Therefore, *grade one* and *these freshmen* form a coreferential relation which is the first type of the identity category. This pair can be used as a good example for teachers to explain the use of co-referential relations between identity devices.

As for example 19, the use of *applicants* and *students* forms the second type of identity relation which refers to the relation between two lexical items expressing different identities to the same referent in the same text. Particularly, *applicants* emphasises one identity of the referents who request to study in a course; while *students* imply another identity of the referents who have been already learning at school. This example provides a good demonstration of using two lexical items to express different identities of the same referent(s), which not only creates the cohesiveness in the text, but also adds more information to the referent(s) in an economic way.

Furthermore, the second point will be discussed with another example. In example 20, the lexical items *English* and *the target language* create an identity pair, in which the former specifies the contextual meaning of the latter. That is to say, *English* is regarded as *the target language* for Chinese learners in this context. This coreferential relation between *English* and *the target language* generate based on this specific context, which exactly demonstrates the highly context-sensitive feature of the identity relations.

Collocation

Collocation has been divided into two sub-categories: activity-related collocation and elaborative collocation in this study. The elaborative collocational pairs have been identified more frequently than the activity-related collocational pairs in the corpus analysis. The reason for this observation may lie in the loose definition of elaborative collocation which only requires two lexical items elaborating or expanding on the same topic, while the activity-related collocation entails the lexical items to be elements of the same activity. Example 21 shows an activity-related collocational relation between *questionnaires* and *participants*. The justification for this interpretation is that in the activity of filling in the questionnaires, *participants* are the ‘people’ who take the action, and *questionnaires* is the ‘thing’ which is the object of the action. That is to say, *participants* and *questionnaires* are two elements (i.e. ‘people’ and ‘thing’) in the ‘activity’, which, therefore, form an activity-related collocational pair.

Example 22 demonstrates the elaborative collocation relations. *Literature* and *research* form an elaborative collocational pair because *literature* triggers the *research* frame, which indicates the occurrence of *research* in the succeeding clause. *Literature* refers to the information relating to the subject *peer feedback*, and *research* denotes the detailed study of the same subject from two different angles which are *the theoretical research* and *students’ perspectives*. The content of *research* provides the information which is included in the

literature. Therefore, both *research* and *literature* are interpreted as elaborating on the topic of research on peer feedback.

It is recommended for the teachers to use them for the teaching of collocation, as these types of collocational relations in these examples are activity-centred or topiccentred, which suggests a greater difficulty of explaining their use out of context only with word meanings of the lexical items.

Conclusion

Compared with studies focusing on misuse of lexical cohesive devices (e.g. Zhang, 2000; Ong, 2011), this study has investigated the appropriate use of lexical cohesion in Chinese postgraduates' writing in UK academic settings, using a tailored framework of lexical cohesion and qualitative analysis of two corpora. The results of this analysis are seen as a useful starting point regarding providing appropriate examples of lexical cohesive devices used in texts for Chinese students' study of such devices, which have hopefully shed some more light on the nature of Chinese students' academic writing regarding their study of lexical cohesion, and the ongoing study of EFL students' writing as a whole.

References

British Council. (2017). *Country brief China April 2007*. Retrieved from siem.britishcouncil.org

Crystal, D. (2008). *A dictionary of linguistics and phonetics*. Oxford: Blackwell Publishing

Flowerdew, J., & Mahlberg, M. (2009). *Lexical and corpus linguistics*. Amsterdam: John Benjamins Publishing.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar*. London & New York: Routledge.

Hoffmann, C. R. (2012). *Cohesive profiling: Meaning and interaction in personal weblogs*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy and other paradigms*. Cambridge: Cambridge University Press.

Oxford English Dictionary Online. (2019). Retrieved from https://www.oed.com/

Ong, J. (2011). Investigating the use of cohesive devices by Chinese EFL learners. *The Asian EFL Journal Quarterly*, *13*(3), 42-65.

Sinicrope, C. (2007). Revisiting cohesive devices in academic L2 English writing : What DO successful writers use ? *Journal of Language Teaching and Research*, *3*(6), 2–25.

Tanskanen, S.-K. (2006). *Collaborating towards coherence*. Amsterdam: John Benjamins Publishing.

Wu, S. (2010). Lexical cohesion in oral English. *Journal of Language Teaching and*

Research, *1*(1), 97–101.

Xi, Y. (2010). Cohesion studies in the past 30 years: Development, application and chaos. *Language Society and Culture*, (31), 139–147.

Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, *31*(1), 61–95.

Zhang, Z. (2018). Academic writing difficulty of Chinese students: The cultural issue behind Chinese and British academic writing styles. *Studies in Literature and Language*, *17*(2), 118–124.

Challenges of Implementing Internationalization of Higher Education in China

Honghuan Li

School of Education, Durham University, Durham, United Kingdom

honghuan.li@durham.ac.uk

Challenges of Implementing Internationalization of Higher Education in China

To achieve the Internationalisation of Higher Education in China, two forms of collaborations were approved by the Ministry of Education in China: the joint university and the joint program. With the apparent success of these joint ventures, more and more universities have joined the team to be ready to operate similar joint ventures in China. Although there are a number of perceived benefits of collaborative universities and programmes in China, several specific challenges exist in transnational education concerning government policy and management, the escalated competition, and education shock faced with running a partnership of collaborative/joint university or program in China in relation to host country and source countries.

Keywords: Internationalization of Higher Education; Transnational Education; Challenges

1. Introduction

Internationalization of Higher Education (IHE) integrates an international, intercultural, or global dimension to cope with the academic environment internationally. The establishment of the formation of international collaborative/joint program in China is defined as Zhongwai-Hezuo-Banxue (中外合作办学in Chinese). The term refers to "activities of the cooperation between foreign educational institutions and Chinese educational institutions in establishing educational institutions within the territory of China to provide education services mainly to Chinese citizens" (MOE, 2003, p. 1). Previous research (Li, 2019) has interpreted the strategies of IHE in China. In this article, the transnational education (TNE) programs development in China, and the challenges in the process of IHE in China will be discussed.

2. Transnational Education Programs in China

According to Knight (2016), the definition of Transnational Education is "the mobility of an education program or higher education institution/provider between countries" (p. 35), which contains "all types of higher education study programs or set of courses of study, or educational services… in which the learners are located in a country different from the one where the awarding institution is based" (UNESCO/Council of Europe, 2002, p. 2). The prime operation of TNE is "cross-border supply" and "commercial presence" as stated in the WTO General Agreement on Trade in Service (Gu, 2009). TNE, more commonly adopted by the providing party, allows higher education institutions to expand themselves in the market and satisfy the demand of the local higher education sector. Not only can it increase the international enrolment by offering their degree qualifications and accreditation in third countries, but can also encourage global participation in higher education (Healey, 2015; McNamara & Knight 2014). In general, the phenomenon of TNE is deemed as the most state-of-the-art stage in the IHE (Doorbar & Bateman 2008; Healey, 2008; Mazzarol et al 2003).

TNE takes various forms, and new forms of TNE are still in the development (Knight, 2016; Burgess & Berquist, 2012); for example, are two styles that will be introduced in the article. International collaborative programs can be various modes, such as international joint/double/multiple degree programs, co-founded or codeveloped institution, locally supported distance education, and twinning program (Knight & Lee, 2012; Knight, 2016). Also, the growth of IBCs is at the leading edge of internationalization by recruiting foreign students and providing education to whom remain in their own home country campuses (Altbach, 2004; Bennell & Pearce, 2003; De Wit, 2002). Collaborative programs and collaborative universities/international branch campuses (IBCs), are the most two popular forms of TNE gaining extraordinary increasing rate development in terms of TNE (Knight, 2016), especially in the process of IHE in China.

In 2003, new regulations of Sino-foreign collaborative/joint schools officially gave the permission of running collaborative programmes and joint universities in China (MOE, 2013). Two forms of collaborations were approved by the Ministry of Education (MOE) in China, the joint universities and the joint programs. The initiation of the collaborations developed relatively slowly at first, surrounded with uncertainties of legislation and regulation (Ennew & Fujia, 2009). However, with the apparent success of these joint ventures and after the joint universities of University of Nottingham Ningbo China (UNNC) and Xi'an Jiaotong Liverpool University (XJTLU) were fully evaluated, more and more universities have joined the team to be ready to operate similar joint ventures in China. Although there are many perceived benefits, concerns and critical aspects of TNE Programs are, however, also explicit from the international community, with the contributors drawing attention to the challenges that foreign providers face operating IHE in China (Helms, 2008).

Shattock (2007) describes the TNE of IHE in China as a tremendous "high-risk experiment". Philip Altbach (2000) even claims a relatively critical opinion on transnational education that

it does not really contribute to the IHE worldwide. Knowledge products are being sold across borders, but with little mutual exchange of ideas, long term scientific collaboration, exchange of students or faculty (Altbach, 2000). There may have cross-cultural influence on teaching and learning in transnational education but one of the pedagogical concerns is "your culture, my classroom, whose pedagogy?" (Crabtree & Sapp, 2004). Other challenges such as degree or certificate accreditation and students' dual identity are worthwhile exploring as well. However, specific challenges regarding government policy and management, competition escalation, and education shock faced with running a partnership of joint universities or programs in China will be discussed below.

3. Challenges

3.1 Government Policy and Management

Since the educational reform needs freedom policies in promoting innovation and creativity, one of the possible problems is that the policy relies dominantly on the host government direction and the host party governance (Paradise, 2012). "The fundamental paradox inherent in this model is that the current reform, which aims to break through excessive administrative control and free up market forces, relies too much on the very political power and authoritarianism that lie behind that administrative control" (Qiang, 2011, page number?). In the partnership, for instance, the operation and the main representative body of the joint program or collaborative university in China should be Chinese. This shows that China is taking its cues and have a relative tendency for keeping up the global education trends by playing a leading role in the partnership of managing collaborative programs or universities. For foreign universities who desire to establish the partnership of TNE programs need to have an agreement on the co-operation or partnership at the level of administering, because the partnership business could encounter the bureaucratically challenging by the governance bodies (Cuiming et al., 2012). Furthermore, managing a joint university in an educational

hub, such as UNNC or XJTLU, needs to liaise closely with the relevant governmental regulatory authorities (Farrugia and Lane 2013) as they operate in a host country in a different cultural context and different ways of running the joint venture. Favourably, China is attempting to decentralize and deduce the bureaucracy in educational matters by offering more power to the provinces and municipality in order to enable higher education institutions to make heavy-handed decisions regarding the educational administration (Jiang, 2011).

3.2 Competition Escalation

Chinese MOE has introduced “Project 211 University” (the 100 leading universities in China) and “Project 985 University” (top universities with the research excellence) to promote the “innovation society” plan in education. Prospective universities are assessed and evaluated by objective and quantifiable criteria on various perspectives, such as buildings, libraries, laboratories, staffing, funds, and research projects. In order to be nominated in the lists, one of the plausible methods is by merging with other universities to expand and promote their academic profile and facilities (Christiansen, 1996; Rosen, 1997). The old Shanghai University, for example, after being merged with other local colleges to come in the new form of Shanghai University, has been successfully selected and nominated as “211 Project University”. Despite of maintaining the internal competition among “Project 211” universities and “Project 985” universities, China is now pursuing the "Double First-Class" initiative to develop its own competitive institutions and disciplines and bring the educational subjects to greater heights in higher education worldwide.

If the “Project 211University” and “Project 985 University” schemes are the internal competition for universities in China, the collaborative universities and programs can be assumed as the external competitors. It is obvious that the IHE in China is experiencing the process of “marketization” (Yin and White, 1994, p. 217) which means that, since the educational suppliers and services are becoming more diversified, the education market will

become more competitive and more intensive with the great increase of consumers' payment ability. The introduction of branch campus from foreign countries may have tremendous impact on local university admissions.

3.3 Education Shock

Similar to culture shock, education shock can be interpreted as the feeling of being confused or anxious when learners or educators have a different educational experience from the one they used to. One of the advantages of TNE programs in China is that students who study in their home countries can earn the qualification awarded by an overseas provider and/or local partner. In the study of Moufahim and Lim (2015), students pointed out that there is a disparity between their expectations of the program and what they experience in the foreign education process. A key issue aroused among Chinese students is that they are seemingly not aware of the educational and cultural diversity of transnational education by equating it to "western education" or "Anglo-Saxon Education". When foreign institutions extend their brand to China, by implementing internationalization of the curriculum and committing to the "internationalization agenda" in TNE programs, the "non-Western cultural issues and topics within courses" (Bennett and Kane, 2009, p. 365) remain rare and "opportunities for truly collaborative and culturally appropriate course design are muted" (Smith, 2010, p. 803). Therefore, it will not be a surprise that students think that they are receiving "Anglo-Saxon" Education by addressing theories, issues, and case studies inside of the western countries. In the TNE programs, the issue of being an international educated person is complicated by the status of acculturation to western education (Chapman and Pyvis, 2006) and there are paradoxes inherent in such experiences (Waters and Leung, 2013).

China in previous years, to some extent, has experienced a certain period of "semi-colonial country" (Huang, 2003, p. 225). This may lead to a distorted view and value that the quality of education from the former colonizer would be more in "superiority" even after the period

of colony. This phenomenon in the process of IHE can also be drawn from the post-colonialism sentiments such as “colonial mentality”, a “psychological construct” held by a people that account for past or present the assumptions, values and beliefs in relation to the colonial power of the postcolonial experiences of colonized peoples (David & Okazaki, 2006, p. 2). The purpose of education turns to the “colonization of the mind”, colonial administration with western forms of knowledge including skills, dispositions and attitudes (Tikly, 2004), or “servitude of the mind” (Altbach, 1977). The education colonialism has also been stated as a form of neo-colonialism in the context of education (Wickens and Sandlin, 2007). The education shock raised by education colonialism, the education post-colonialism or the education neo-colonialism, to some extent, provides a crucial supplement to the analysis of the challenges.

One of the characteristics of running a TNE program in China is that the collaborative programs should be partially delivered by the partnering institutions in China. In that circumstance, a teacher from a Chinese institution may have to refrain from using their own cultural pedagogy and will rather adapt to the western pedagogy as is required by the foreign partner, and presumably the western pedagogy model is more sophisticated and innovative than their own model. The joint universities functioning as international branches, students from oversea countries can also benefit from Chinese culture and pedagogy, not merely Chinese students. To reduce of the shock effect, UNNC, for example, has integrated courses relating to Chinese culture to the curriculum for better understanding of the differences and reaching the consensus on the importance of both Chinese and western cultural and the educational communication in between.

4. Conclusion and Implications

In summary, this article has briefly introduced the context of Internationalisation of Higher Education and the working mechanisms of the transnational education programs in China.

Although it seems to be a risky experiment for China to implement IHE, the TNE programs provide more opportunities, not only greater chances for students to access educational resources worldwide, but also for better and faster development for the country in various aspects. This article also pointed out the challenges of running joint ventures in China, respectively, government policy and management, competition escalation, and education shock. While this challenges ahead are unavoidable, China is still finding its way to resolve these issues, such as decentralizing and deducing the bureaucracy in educational administration, promoting and constructing "Double First-Class" universities worldwide, integrating Chinese cultural and educational elements into TNE programs etc. Further research could be focused on the effectiveness of the TNE program as well as the future development and action plans of the IHE process in China.

References

Altbach, P. G. (1977). Servitude of the Mind? Education, Dependency, and Neocolonialism. *Teachers College Record*, 79(2), 187-204.

Altbach, P. G. (2000). The crisis in multinational higher education. *Change: The Magazine of Higher Learning*, 32(6), 28-31.

Altbach, P. G. (2004). Higher education crosses borders: Can the United States remain the top destination for foreign students?. *Change: the magazine of higher learning*, 36(2), 18-25.

Bennell, P., & Pearce, T. (2003). The internationalisation of higher education: exporting education to developing and transitional economies. *International Journal of educational development*, 23(2), 215-232.

Bennett, R., & Kane, S. (2009). Employer engagement practices of UK business schools and departments: an empirical investigation. *Journal of Vocational Education and Training*, 61(4), 495-516.

Burgess, P. and Berquist, B., 2012. Cross-border delivery: programs, programs, and providers. In Deardorff, D., de Wit, H., Heyl, J. and Adams, T. (eds.) *The SAGE handbook of international higher education*. London: SAGE, pp325-342.

Chapman, A., & Pyvis, D. (2006). Dilemmas in the formation of student identity in offshore higher education: A case study in Hong Kong. *Educational Review*, 58(3), 291-302.

Christiansen, F. (1996). *Devolution in Chinese Higher Education Policy in the 1990's: Common Establishment and the" 211" Programme*. Department of East Asian Studies, University of Leeds.

Crabtree, R. D., & Sapp, D. A. (2004). Your culture, my classroom, whose pedagogy? Negotiating effective teaching and learning in Brazil. *Journal of Studies in International Education*, 8(1), 105-132.

Cuiming, G., Feng, Y., & Henderson, F. (2012). On joint-programs in China– development, challenges and suggestions. *On the Horizon*, 20(4), 293-304.

David, E. J. R., & Okazaki, S. (2006). Colonial mentality: A review and recommendation for Filipino American psychology. *Cultural Diversity and Ethnic Minority Psychology*, 12(1), 1.

De Wit, H. (2009). *Internationalization of higher education in the United States of America and Europe*. IAP.

Doorbar, A., & Bateman, C. (2008). The growth of transnational higher education: The UK perspective. *Teaching in transnational higher education: Enhancing learning for offshore international students*, 14-22.

Ennew, C. T., & Fujia, Y. (2009). Foreign Universities in China: a case study. *European Journal of Education*, 44(1), 21-36.

Farrugia, C. A., & Lane, J. E. (2013). Legitimacy in cross-border higher education: Identifying stakeholders of international branch campuses. *Journal of Studies in International Education*, 17(4), 414-432.

Healey, N. (2015). Managing international branch campuses: what do we know?. *Higher Education Quarterly*, 69(4), 386-409.

Healey, N. M. (2008). Is higher education in really 'internationalising'?. *Higher Education*, 55(3), 333-355.

Helms, R. M. (2008). *Transnational education in China: Key challenges, critical issues and strategies for success*. Observatory on Borderless Higher Education.

Huang, F. (2003). Policy and practice of the internationalization of higher education in China. *Journal of Studies in International Education*, 7(3), 225-240.

Jiang, K. (2011). De-bureaucratization within China's Universities. *Inside Higher Education*, 12.

Jianxin, G. U. (2009). Transnational education: Current developments and policy implications. *Frontiers of Education in China*, 4(4), 624-649.

Knight, J. (2016). Transnational education remodeled: Toward a common TNE framework and definitions. *Journal of Studies in International Education*, 20(1), 34-47.

Knight, J., & Lee, J. (2012). International Joint, Double, and Consecutive Degree Programs. *The Sage Handbook of International Higher Education*, 343.

Li, H. (2019). *Internationalisation of higher education strategies in China: using institutional theory and demand-driven analysis*. Durham University, School of Education. http://dro.dur.ac.uk/27669/1/27669.pdf

Mazzarol, T., Norman Soutar, G., & Sim Yaw Seng, M. (2003). The third wave: Future trends in international education. *International Journal of Educational Management*, 17(3), 90-99.

McNamara, J., & Knight, J. (2014). *Impacts of Transnational Education on Host Countries: Academic, Cultural, Economic and Skills Impacts and Implications of Programme*

and Provider Mobility [Internet]. British Council/DAAD. London: British Council and DAAD.

MOE. (2003). Retrieved February 07, 2017, from http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_861/200506/8644.html

Moufahim, M., & Lim, M. (2015). The other voices of international higher education: an empirical study of students' perceptions of British university education in China. *Globalisation, Societies and Education*, 13(4), 437-454.

Onsman, A. (2011). It is better to light a candle than to ban the darkness: Government led academic development in Saudi Arabian universities. *Higher Education*, 62(4), 519-532.

Paradise, J. (2012). International Education: Diplomacy in China. *The Brown Journal of World Affairs*, 19, 195.

Qiang, Z. (2003). Internationalization of higher education: Towards a conceptual framework. *Policy Futures in Education*, 1(2), 248-270.

Qiping, Y., & White, G. (1994). The 'marketisation' of Chinese higher education: A critical assessment. *Comparative Education*, 30(3), 217-221.

Rosen S. 1997. Recent developments in Chinese education. *Chinese Education and Society*, 30(3); 1-9.

Shattock, M. (2007). *Overseas campuses: the management perspective. British Universities in China: The Reality Beyond the Rhetoric*. An Agora Discussion Paper, December http://www. agora-education.

Smith, K. (2010). Assuring quality in transnational higher education: a matter of collaboration or control?. *Studies in Higher Education*, 35(7), 793-806.

Tikly, L. (2004). Education and the new imperialism. *Comparative education*, 40(2), 173-198.

UNESCO/Council of Europe. (2002). *Code of good practice in the provision of transnational education*. Retrieved from http://www.coe.int/t/dg4/highereducation/recognition/code of good practice_EN.asp

Waters, J., & Leung, M. (2013). A colourful university life? Transnational higher education and the spatial dimensions of institutional social capital in Hong Kong. *Population, Space and Place*, 19(2), 155-167.

Wickens, C. M., & Sandlin, J. A. (2007). Literacy for what? Literacy for whom? The politics of literacy education and neocolonialism in UNESCO-and World Bank– sponsored literacy programs. *Adult Education Quarterly*, 57(4), 275-292.

(Re)constructing imagined identities in language cafés: an ethnographic inquiry

Nuria Polo-Pérez

School of Education, Durham University, Durham, United Kingdom

nuria.polo-perez@durham.ac.uk

Nuria Polo-Pérez is an Assistant Teaching Fellow in Spanish at the Centre for Foreign Language Study and a Doctoral student in the School of Education at Durham University.

(Re)constructing imagined identities in language cafés: an ethnographic inquiry

Despite their growing popularity, language cafés have not received the same research attention as other second or foreign language contexts outside of formal education. Drawing on ethnographic data from a three-year fieldwork conducted in a university language café, this paper aims to draw attention to language cafés as significant social contexts for multilingual identity performance and development. In particular, underpinned by a poststructuralist view of identity, the findings show different ways in which participants (re)construct their imagined multilingual identities in interaction with others in the language café environment, where subject positions are not constrained by institutional roles (e.g. student and teacher).

Keywords: language cafés; language socialisation; imagined identities.

Introduction

'Language café' (LC) is used here as an umbrella term to designate any group that provides non-formal regular events for individuals to practise their foreign or second language(s) (LX[1]) by interacting with others. They are non-formal because they do not provide any formal instruction as such, and they do not belong to any institutional programme. The events are generally organised by volunteers and in public spaces, such as bars, libraries, cafés, etc. Despite the opportunities for meaningful LX socialisation that these events seem to foster, LCs have not received the same research attention as other learning contexts outside of formal instruction. This is even more so with regards to contexts where English is not the target language.

This paper addresses one of the emerging themes from a doctoral research project which investigates ethnographically and from an ecological perspective (van Lier, 2004) the affordances for multilingual and intercultural identity development in a LC in North England (UK). In particular, by looking at how learners make sense of their experiences in the LC, this paper will shed light on the ways in which the LC affords the (re)construction of learners' imagined (multilingual) identities.

[1] The use of 'LX' instead of 'L2', and 'L1 user' instead of 'native speaker', follows the recommendations by Jean-Marc Dewaele (2018) in his article entitled 'Why the Dichotomy 'L1 Versus LX User' is Better than 'Native Versus Non-native Speaker'.

An ecological approach to language learning

The way I approach the study of language socialisation in the non-instructional setting of LCs draws on the 'ecology' metaphor, 'which captures the dynamic interaction between language users and the environment as between parts of a living organism' (Kramsch, 2002, p. 3). Thus, an ecological approach sees L2 development as a result of meaningful participation in human events (van Lier, 2004); that is, a complex, non-linear, dynamic and emergent process that is based on the interactions and interrelations between the learners and their ecosystem (Kramsch, 2002; Kramsch & Whiteside, 2008; van Lier, 2004). An ecological perspective, therefore, avoids simple cause-effect relationships and aims at exploring the interconnectedness of different elements in the ecosystem. With this approach, I also embrace the perspective that there is no separation between language acquisition and language socialisation.

Language cafés as a complex and under-researched context

While everyone has an idea of what the ecosystem of a classroom may look like, people refer to very different realities when they use the term 'language café' (or any of its equivalent denominations). LCs are too diverse to be portrayed as a homogeneous type of event, since they often differ in the key dimensions that, according to Benson (2011), ought to be considered when researching language learning beyond the classroom:

- location: where the activity takes place;
- formality: the degree to which the learning is aligned with an institutional programme;
- pedagogy: the extent to which teaching practices are involved; and
- locus of control: the extent to which learners self-manage their environment.

In this regard, in some LCs the organisers take up the role of teacher-facilitators during the conversations (and may even charge for it) and, on the other side of the spectrum, in other LCs, all participants share an equal status in the management and negotiation of their interactions. The LCs involved in my study belong to the latter group. Nevertheless, these dimensions should be considered as fluid and negotiable processes, rather than static settings in the environment. Furthermore, in the case of LCs, another important dimension to consider might be how stable or transient these groups are, since this interferes with the affordances these environments may be able to provide.

Previous research on different informal language learning meetups has mostly been limited to contexts of English as a foreign language (EFL), unlike the UK-based LCs in my study, where people gather to speak different languages other than English. Due to the socioeconomic status of English as a global language, some scholars have argued that research findings from EFL contexts might not always apply to other language contexts, for instance, when it comes to learning motivations (Duff, 2017).

In Turkey, Balçıkanlı (2017) shows how the English Café is perceived as a place (1) to practise English, (2) to socialise, (3) to exchange knowledge and life experiences, (4) to learn from others and (5) a safe place to take risks and deal with feelings and emotions. In Japan, Murray and Fujishima (2016) gather narratives from different stakeholders in the L-café, a university facility which provides a space for Japanese and international students to practice their languages (mainly English) informally. In a later study, they focus on aspects of participant entry, access, and belonging in this same social learning environment (Murray, Fujishima, & Uzuka, 2017). These recent studies, however, feature LCs where teacher-facilitators intervene to enhance the learning experience, unlike the LC in my study, where the groupings and conversations are completely spontaneous and self-managed by participants.

In China, Gao (2009) studied participant experiences in an English Club through the thematic analysis of their spontaneous webforum 'reflective experiential accounts'. These accounts revealed perceptions of the English club as a place for supportive peer-learning, self-assertion and changes in self-perception. Gao (2009), thus, reclaims a more humanistic approach to language teaching whereby social relationships are promoted, so as to help learners sustain their autonomous learning efforts. Building up on Gao's qualitative work, Liu (2013) uses statistical data to study the development of students' self-efficacy as a result of participating in a Chinese university English Bar. Self-efficacy is defined in that context as 'students' perceived capability to conduct effective verbal communications with native and non-native speakers' of the target language (Liu, 2013, p. 28). However, I will argue that interacting in LCs goes beyond engaging in effective communication. Participating in LCs is about LX socialisation more broadly and, as such, it involves the embodied experience of performing one's full social being in another language, and not just one's student self.

Imagined identities and language learning

This paper draws on a poststructuralist understanding of identities as multiple and dynamic; in other words, identities are subject to a life-long process of non-linear transformation, and they are always complex, socially constructed, and historically situated (Kramsch, 2009; Norton, 2013; Norton-Peirce, 1995; Pavlenko & Lantolf, 2000; van Lier, 2004). In line with a social constructionist approach, this study assumes that people construct and re-construct their sense of self in interaction with others. Through social interaction people position themselves, and are positioned by others, in relation to the social world around them (Kramsch, 2009).

With an encompassing view of multilingualism, any individual who is engaged in additional language learning or usage can be considered as a 'multilingual speaker', regardless of the number of languages they speak or their level of competence (Fisher, Evans, Forbes, Gayton, & Liu, 2018). The importance of developing a multilingual identity has been associated with the positive impact it has over learners' effort investment in learning and maintaining their languages, and with its potential to enhance social cohesion in nowadays increasingly diverse

societies (Fisher et al., 2018). Although the classroom can be an effective first site for learners to come into contact with their subjectivities as multilingual speakers (Kramsch, 2009), it is when learners start living the LX from within, rather than studying it as an objective reality that exists at a distance, that they start seeing themselves as *languagers* and full social beings in the material worlds that surround that language (Phipps & Gonzalez, 2004).

Furthermore, the importance of projecting oneself in the future has been recognised as central to the lives of many LX learners, since imagined or desired identities also have an impact on individuals' motivation to learn and, most importantly, in how invested learners are in the language (Norton, 2013). According to Norton, motivation, like identity, cannot be seen as a static feature in the learner. Moreover, a highly motivated learner might be unsuccessful because of the unequal power relations they encounter in communication with target language speakers. Norton (2000) proposes the concept of 'investment' as a sociological alternative to the psychological construct of 'motivation'. As she puts it,

> if learners "invest" in the target language, they do so with the understanding that they will acquire a wider range of symbolic resources (language, education, friendship) and material resources (capital goods, real estate, money), which will in turn increase the value of their cultural capital and social power. As the value of language learners' cultural capital increases, so learners' sense of themselves, their hopes for the future, and their imagined identities are reassessed. Hence there is an integral relationship between investment and identity, an identity which is theorized as multiple, changing, and a site of struggle. (Norton, 2013)

Considering all of the above, the main question addressed in this paper is: How do face-to-face interactions in the LC contribute to participants' (re)construction of selves and their imagined identities?

Methods

My research contributes to the groundbreaking literature on LCs by studying these environments from a complex researcher positioning. Rather than being a detached observer, I participated in the LC events as a genuine LX user and learner, drawing on my multilingual self as a resource in all the stages of the ethnographic research (Holmes, Fay, Andrews, & Attia, 2013, 2016).

I collected ethnographic data over a period of three years in two different LCs, where I actively participated as a multilingual speaker, with a particular interest in improving my French speaking skills. I kept a record of my experiences in a reflective journal which reached approximately 35,000 words; I audio-recorded around seven hours of naturally-occurring conversations in the LCs, and conducted 16 semi-structured interviews with participants, eight of whom also completed short written reflections about their experiences in the LC. For the purpose of further data triangulation, I also gathered fieldnotes based on my observations of two LC sessions where I participated as a complete observer. Audio-

recorded data were transcribed and analysed using thematic analysis (Braun & Clarke, 2006). All informants have pseudonyms to maintain their anonymity. In order to provide a clear focus, in this paper I will draw on the interview data from only one of the two LCs I investigated, namely the university LC.

This university LC is located in North England and is organised by university teachers. It takes places twice per term at the student union bar, and thus attracts mainly university students. As any extracurricular activity at the university, the LC is not attached to any study programme and attendance is voluntary. This LC is multilingual, in that it is aimed at practising different languages during the same evening. An important characteristic of this LC is that most conversation groups during the events are self-managed by participants without the intervention of any designated group facilitators. Groups tend to include speakers with mixed abilities in the target language(s), which affords the emergence of pedagogic episodes of collaborative peer learning. This diversity, I will argue, also affords meeting inspiring others and the reconstruction of multilingual imagined identities.

(Re)constructing one's imagined identities in the LC

The findings show that participants reconstruct their imagined identities in LCs as a result of meaningful socialisation experiences in their LX. This is connected to the fact that these experiences often trigger a change in learners' self-perception as multilingual speakers, or prompt their self-projection as successful multilinguals inspired by more competent others. In both cases, feelings of reassurance support their imagining of new and encouraging possibilities for their future selves.

Imagined identities through changes in self-perception

Participants in this university LC conceptualise the environment and their experiences in it around a main underlying idea: a sense of freedom. Because most of these participants are students, the classroom seems to be the most immediate context with which they tend to compare the LC environment. Thus, they frequently remark that people in the LC are free to stay as long they want, talk about anything they want, and make mistakes without the judgment of a teacher.

Yet freedom also involves that, as opposed to the classroom, the LC represents an unguided and unpredictable social space, which might be daunting for many first-timers. This is why self-doubt before attending the LC (e.g.: Is my level good enough? Will I fit in?) is very frequent. However, by the end of the event, self-doubt sentiments tend to transform into a strong sense of satisfaction and achievement (e.g.: I did it! I have spoken for over an hour in [language]!), which seems to be directly related with self-assurance and motivation to take up on new challenges in the future, as illustrated by the two interview comments below:

> [This sense of satisfaction] makes you want to do more, and makes you think like "oh, next time I'm gonna do this and it's gonna be…" or "next time I'm gonna be even better!" (Amy)
> When somebody speaks really well and has a really good [accent]… Well, actually, I don't think it comes from listening to other people that are better, I think it comes from just enjoying it and wanting to do better next time. […] So it's like… happy with the things you've done… it's like 'I know I've made a lot of mistakes too, but next time I don't wanna do that [laughter]'. (Rebecca)

For Amy, the LC experience has set a precedent of successfully managing an informal social event in Spanish for the first time. Having this precedent in mind, she is now able to imagine a future self taking up new challenges and performing even better. Those who have experienced a sports achievement, such as running 10K for the first time, will be familiar with this type of post-event motivation boost. However powerful, it is important to acknowledge that this type of self-assurance can be also very ephemeral and does not necessarily translate into long-term, sustained motivation.

Rebecca, on the other hand, is an invested polyglot and regular attendee whose experiences in the LC are always related to reconnecting with the joy of speaking languages, which keeps her imagination as a future fluent LX user alive.

Imagined identities through self-projection in others

LC participants have a tendency to tacitly map their interlocutors' approximate levels of proficiency during the first minutes of interaction with others in the LC. By comparing with each other, they are constantly reassessing and adjusting their positionings, self-perception and self-projection as multilinguals. In relation to this, and slightly related to learning motivation, 'jealousy' is frequently mentioned in interview comments as a feeling that emerges in interaction with more competent others in the LC:

> Whenever they [other LC participants] are actually doing a language degree and they talk better than me, I accept it, but I'm very jealous at the same time, 'cos I'm like 'I wish I could talk like that!' (Elisabeth)

What Elisabeth implies is that with time and investment she believes she could be as fluent as those successful speakers she meets in the LC. Interacting with advanced LX users can arguably support the construction of her imagined identities based on attainable goals, which move away from the unattainable model of the native speaker.

On the other hand, fluent speakers in the LC can also see themselves projected retrospectively in those who are on the first stages of their learning. In the following excerpt,

Nathan recalls a conversation with a beginner student, and how he could identify with his puzzlement when interacting with an advanced LX user:

> (…) cuando le dije que estaba en Stage 6, me vio como… ¡pues eso es otro mundo, sabes! (…) En francés antes cuando yo no tenía tanto nivel, y yo veía a gente que tenía… no sé, que hablaba muy bien, aunque no fuera muy bien, ¡ya para mí era increíble! (Nathan)
> *(…) when I told him that I was in Stage 6, he looked at me like… that's a different world, you know! (…) In French, when I was not that fluent and I saw somebody with… I don't know, speaking very well, even if it was not that well, for me that was already incredible! (Nathan)*

Nathan then goes on to explain that, just like others inspired him in the past, he really enjoys now being a source of inspiration and motivation for others in the LC. He shows empathy with those who are starting their learning paths as a way to be reciprocal retrospectively.

Other examples of inspiring others often emerge in interaction with older and more experienced LX users. When asked about a memorable moment from an LC, Molly recalls a conversation from her hometown LC with a woman who successfully embedded multilingual practices in her professional career:

> (…) she was just saying how much she enjoyed being able to speak both her languages in the office, and it kind of just introduced the idea of being able to… maybe I could work in the UK and in France. (Molly)

At the time of that conversation, Molly was about to start her modern languages degree at university. This woman had a powerful, inspirational effect on Molly's construction of her imagined identities as a future professional with the rewarding experiences of working in multilingual environments and dwelling in languages.

Discussion

The findings show that interacting with others in LCs does not only involve information exchange, but participants are constantly reorganising a sense of who they are and their desires for the future; that is, through participation, they engage in the construction and reconstruction of selves and their imagined identities (Norton, 2013; Norton-Peirce, 1995; Pavlenko & Lantolf, 2000). Some participants, like Molly, feel inspired by others to imagine themselves as cosmopolitan speakers with a 'nomadic and borderless lifestyle' (Ros i Solé, 2013, p. 326). The reconstruction of selves and the development of imagined identities as multilinguals are, therefore, intersubjective processes that are not anchored in the individual's mind, but embedded in the sociality of LX learning (Ros i Solé, 2016). These processes are also afforded by the ecologies of the LC as a peer-learning environment, where reciprocity and empathy provide the power-balanced ground that some learners need to feel free to exercise their agency as LX users.

Rather than just practising their LX, the findings suggest that LC participants engage in *languaging*, understood as the 'full, embodied and engaged interaction with the world that comes when we put the languages we are learning into action' and 'the effort of *being* a person in that language in the social and material world of everyday interactions' (Phipps, 2007, p. 12, my emphasis). Whereas we cannot underestimate the value of systematic learning in the classroom (Woodin, 2018), the ecologies of the LC environment afford the full embodied experience of informal language socialisation and the emotions associated with it. Rewarding social experiences in the target language mobilise a new sense of self as a multilingual speaker in a way that goes far beyond the satisfaction that one might get from a good grade. Thus, as language educators we should ask ourselves: How does my programme support students' experiences of *languaging*?

Finally, Risager (2006) points to the discussion of how we 'ought to upgrade and include non-native speakers as linguistic models in language teaching and language practice' (p. 130). The term 'native speaker' has been widely critiqued for its racist connotations (Holliday, 2006), and for revealing a monolingual bias and perpetuating a deficit model in language learning, whereby non-native speakers are defined by what they did not achieve (Cook, 1999; Dewaele, 2018). Byram (1997) argues that the native speaker model sets an impossible target for learners and evaluates the wrong kind of competence; instead, language learners should aim at becoming intercultural speakers. The findings in my study suggest that, in the LCs, successful LX users can act as role models and a powerful source of inspiration for learners, as they support the reconstruction of learners' imagined future identities based on achievable goals that move away from the unattainable model of the native speaker.

Conclusions

This paper reported on one of the salient themes emerging from an ethnographic inquiry into the affordances for languaging and the development of multilingual identities in a multilingual LC in North England. In particular, it explored the ways in which LC participants reconstruct their imagined identities and their sense of self as multilinguals as a result of their socialisation experiences in LCs. Thanks to the particular ecologies of these environments, characterised by the absence of the traditional teacher and student roles, and supported by a shared, strong sense of reciprocity and empathy, LC participants often undergo a change in self-perception as LX users as they experience the enjoyment of how it feels to socialise and create new human connections in another language (Phipps & Gonzalez, 2004). These new human connections, in turn, are often a great source of inspiration which prompts participants' self-projection of imagined identities as successful multilingual and cosmopolitan speakers. These changes in self-perception and self-projection in others encompass feelings of reassurance which support their LX learning efforts and investment, as they imagine new and encouraging possibilities for their future multilingual selves (Norton, 2013).

References

Balçıkanlı, C. (2017). The 'English Café' as a social learning space. In G. Murray & T. Lamb (Eds.), *Space, Place and Autonomy in Language Learning* (pp. 61-75). London: Routledge.

Benson, P. (2011). Language learning and teaching beyond the classroom: An introduction to the field. In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 7-16). Basingstoke: Palgrave Macmillan.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters.

Cook, V. (1999). *Going beyond the native speaker in language teaching*. Washington, D.C.: TESOL.

Dewaele, J.-M. (2018). Why the Dichotomy ' L1 Versus LX User' is Better than 'Native Versus Non-native Speaker'. *Applied Linguistics, 39*(2), 236-240. doi:10.1093/applin/amw055

Duff, P. A. (2017). Commentary: Motivation for Learning Languages Other Than English in an English- Dominant World. *Modern Language Journal, 101*(3), 597-607. doi:10.1111/modl.12416

Fisher, L., Evans, M. J., Forbes, K., Gayton, A., & Liu, Y. (2018). Participative multilingual identity construction in the languages classroom: a multi-theoretical conceptualisation. In: Apollo - University of Cambridge Repository.

Gao, X. (2009). The ' English corner' as an out-of-class learning activity. *ELT Journal, 63*(1), 60-67. doi:10.1093/elt/ccn013

Holliday, A. (2006). Native- speakerism. *ELT Journal, 60*(4), 385-387. doi:10.1093/elt/ccl030

Holmes, P., Fay, R., Andrews, J., & Attia, M. (2013). Researching multilingually : new theoretical and methodological directions. *International journal of applied linguistics, 2013, Vol.23(3), pp.285-299.*

Holmes, P., Fay, R., Andrews, J., & Attia, M. (2016). How to research multilingually: Possibilities and Complexities. In H. Zhu (Ed.), *Research Methods in Intercultural Communication: A Practical Guide* (pp. 88-102). New York: Wiley Blackwell.

Kramsch, C. (2002). *Language acquisition and language socialization : ecological perspectives*. London :: Continuum.

Kramsch, C. (2009). *The multilingual subject : what foreign language learners say about their experience and why it matters*. Oxford :: Oxford University Press.

Kramsch, C., & Whiteside, A. (2008). Language Ecology in Multilingual Settings. Towards a Theory of Symbolic Competence. *Applied Linguistics, 29*(4), 645-671. doi:10.1093/applin/amn022

Liu, M. (2013). English Bar as a Venue to Boost Students' Speaking Self-Efficacy at the Tertiary Level. *English Language Teaching, 6*(12), 27. doi:10.5539/elt.v6n12p27

Murray, G., & Fujishima, N. (2016). *Social spaces for language learning : stories from the L-café*. Basingstoke, Hampshire: Palgrave Macmillan.

Murray, G., Fujishima, N., & Uzuka, M. (2017). Social learning spaces and the invisible fence. In G. Murray & T. Lamb (Eds.), *Space, place and autonomy in language learning* (pp. 233-245). London, New York: Routledge.

Norton, B. (2000). *Identity and Language Learning: Gender, Ethnicity and Educational Change*. Harlow: Longman.

Norton, B. (2013). *Identity and Language Learning: Extending the Conversation*. Bristol: Multilingual Matters.

Norton-Peirce, B. (1995). Social Identity, Investment, and Language Learning. *TESOL Quarterly, 29*(1), 9-31. doi:10.2307/3587803

Pavlenko, A., & Lantolf, J. P. (2000). Second language learning as participation and the (re)construction of selves. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 155-177). Oxford: Oxford University Press.

Phipps, A. M. (2007). *Learning the arts of linguistic survival : languaging, tourism, life*. Clevedon :: Channel View.

Phipps, A. M., & Gonzalez, M. (2004). *Modern languages : learning and teaching in an intercultural field*. London :: Sage.

Risager, K. (2006). *Language and culture : global flows and local complexity*. Clevedon: Multilingual Matters.

Ros i Solé, C. (2013). Cosmopolitan speakers and their cultural cartographies. *The Language Learning Journal, 41*(3), 326-339. doi:10.1080/09571736.2013.836349

Ros i Solé, C. (2016). *The personal world of the language learner*: London : Palgrave Macmillan.

van Lier, L. (2004). *The ecology and semiotics of language learning : a sociocultural perspective*. London: Kluwer Academic Publishers.

Woodin, J. (2018). *Interculturality, interaction and language learning: insights from tandem partnerships*. London: Routledge.

English for Academic Purposes (EAP) in UK higher education: examining the impact of EAP's position within the academy on service delivery, identity and quality

Sharon Riddle

School of Education, Durham University, Durham, United Kingdom

sharon.l.riddle@durham.ac.uk

Sharon Riddle's research focuses on the positioning of English for Academic Purposes in UK higher education.

English for Academic Purposes (EAP) in UK higher education: examining the impact of EAP's position within the academy on service delivery, identity and quality

This paper will outline a forthcoming PhD project exploring the provision of English for Academic Purposes (EAP) services in UK higher education institutions. Neoliberalism in UK universities has contributed to the extent and importance of EAP, but has also led to many challenges for staff and service delivery, including where EAP is located in HE structures. The study will examine EAP's position, focusing on questions relating to its role, how EAP practitioners operate, the impact of the position of EAP, and improvement strategies. Three EAP units in different structural positions in higher education will be examined and data will be considered in the light of Bourdieu's theories of 'field', 'habitus', 'capital' and 'symbolic violence'. It is hoped that the study will shed light on why EAP services are positioned as they are, how this impacts on quality, and how they can reposition to provide more benefit to the academy.

Keywords: academic literacy; English for Academic Purposes; international student education; higher education

The neoliberalisation of higher education has transformed universities into commercial enterprises which have to grapple with tensions between business imperatives and academic standards (Murray, 2015). Marketization and changes in the financing of higher education have forced all universities to secure alternative funding, leading to increased competition in the international student market (Foskett, 2011). This has led to a rapid rise in high fee-paying international students in the UK, with a total of 458,490 in 2017/18 (HESA, 2019). Success in attracting and retaining international students is key, and involves providing value for money, a positive student experience and ultimate academic success, resulting in enhanced income and prestige for the institution. Academic literacy is a major issue for international students, leading to the increased importance of language support, usually known as English for Academic Purposes (EAP).

The provision of high quality language support plays a key part in the internationalisation agenda and financial health of universities. The marketization of higher education and the resulting rapid rise in international students has increased the demand for EAP, shaped its role and has led to a shift in its position in universities. While it is recognised in the literature that EAP is 'on the edge' of academia (Ding & Bruce, 2017), there has been no detailed work on causes, effects and solutions. This article describes a forthcoming study which aims to fill this gap by examining the position of EAP within UK universities, and the impact on identity, quality and service delivery.

There are many definitions of EAP, including EAP being "famously needs-driven, centrally focused on meeting the needs of students seeking to study or undertake research in English-medium university contexts" (Bruce, 2017, p. 1). However, these definitions fail to represent the complexity of current provision (Campion, 2016). The complex landscape of EAP

provision (Blaj-Ward, 2014) includes foundation, pre-sessional and in-sessional courses, provision of language support for lower level students (Alexander, 2012; Murray, 2015), decentralised services to engage more effectively with academic departments (Murray, 2016) and outsourced services in the private sector.

Neoliberalism has had a significant impact on the role, governance and funding of UK higher education, including how students are viewed and the nature of academic life and identity. The marketization of higher education has transformed universities into commercial enterprises (Foskett, 2011) which compete for student consumers (Murray, 2015), leading to a shift in the position and role of EAP to that of servicing high numbers of fee-paying students, for example on large pre-sessional programmes. Therefore, EAP courses attract substantial income for universities, but they are also often under-resourced and subject to cost-cutting measures (Marginson, 2010). This has resulted in the perception of EAP as an income generator at a sub-degree level with reduced academic standing (Fulcher, 2009). Also, universities are affected by increased privatisation and outsourcing (Ball, 2007), particularly in the provision of academic literacy services, resulting in private EAP providers working with 61 British universities in 2016 (Bell, 2016). In addition, the process of managerialism has resulted in changed roles, an imitation of private sector management practices including an increase in precarious short-term contracts (Deem, Hillyard, & Reed, 2007; Hyland, 2018) and an erosion of the professional autonomy of academics (Olssen & Peters, 2005). New roles have emerged, for example ‘third space’ professionals who span professional and academic domains (Whitchurch, 2013).

The existence and growth of EAP is rooted in the transformation of higher education (Ding, 2016) together with the spread of English (Ferguson, 2007; Phillipson, 2009). However, these

same factors have also influenced the positioning of EAP within universities, resulting in it occupying different spaces within organisational structures, including units which are on the margins (Ding & Bruce, 2017), outsourced (Fulcher, 2009) or decentralised (Murray, 2016; Wingate, 2018). A direct result of neoliberalism has been the development of a 'third space' professionals who work across the academic and professional divide (Whitchurch, 2013). In the EAP context, Hadley (2015) studied Blended EAP Professionals (BLEAPs), whose role is a mixture of teaching and management, working in 'Student Processing Units' which provide "academic training, remedial learning, and other educational experiences for large numbers of students" (p. 39). Furthermore, Hyland and Hamp-Lyons (2007) note that in these changed institutional contexts, "EAP teachers are frequently employed as vulnerable, short-term instructors in marginalized 'service units'" (2007, p. 10), and this theme of marginalisation is touched on repeatedly in the literature (Ding & Bruce, 2017; Turner, 2004).

Marginalisation is usually discussed in relation to social exclusion and a sense of not belonging (Mowat, 2015), leading to powerlessness (Hyland, 2018) and oppression (Moran, 2014). In the context of EAP provision in UK higher education, it is manifested in a number of ways. Organisational structures increasingly place EAP services outside of the university, resulting in the proliferation of private providers (Bell, 2016; Fulcher, 2009). The lack of status and influence of EAP practitioners is also affected, where practitioners are vulnerable, professionally isolated or powerless with limited job security and little influence (Ding & Bruce, 2017; Hadley, 2015; Hyland, 2018). Furthermore, EAP can be invisible in the literature related to the internationalisation of universities, for example, De Vita (2007), suggesting that there is no guarantee that EAP will be involved in planning for international students.

Differing perceptions and a lack of understanding of EAP has also weakened its position, leading to EAP practitioners feeling undervalued (Murray, 2016) with a lack of recognition of EAP's wider role in enhancing disciplinary writing and providing academic literacy support. In fact, some believe that EAP is guilty of colluding in its own marginalisation by taking a 'butler's stance' (Raimes, 1991) and accepting a role as an "economic and intellectual shortcut" (Turner, 2004, p. 86). Furthermore, with no common prerequisite qualification for becoming an EAP practitioner (Ding & Campion, 2016), academic colleagues may believe that EAP does not require any specialist skill and can be taught by anyone who can speak English (Murray, 2016).

Mowat (2015) advocates that marginalisation encompasses both the state of being marginalised, and also the resulting emotional response, including feelings about not belonging and frustrations about not being able to work effectively and contribute. Hadley's findings agreed with this, where he found that some EAP units are places of fear, anger, disappointment and "psychological suffering" (Hadley, 2015 p. 359).

In responding to marginalisation, Johns suggests that EAP practitioners should be proactive and not simply "wring hands and bemoan marginality" (1997, p. 154). Practitioners are also urged to take a critical stance by not accepting the status quo, but question the political context of EAP (Benesch, 2001). In the context of the strong structural forces of neoliberalism, Ding and Bruce (2017) recommend adopting a reflexive attitude to develop agency, and that groups, or 'corporate agents', can exercise power using collective action (Archer, 2003), a possible role for BALEAP, EAP's professional body.

The positioning of EAP within the academy has had an effect on the identity of practitioners. A sense of identity is formed through combining internal and external views of an individual or group (Jenkins, 2008), and there is sometimes a contrast between these views, as in the case of EAP. The 'inside view' is that EAP is key in supporting students and researchers whose academic success relies on them being able to communicate in English (Hyland, 2018). Benesch (2001) takes this further, suggesting that EAP is transformative, helping students to not only do well, but also encouraging them to question and shape the education they are being offered. Also, EAP is also considered to be an academic discipline and research activity (Ding & Bruce, 2017; Hamp-Lyons, 2011). However, there are also issues regarding the identity of EAP and EAP practitioners, with EAP seen as being on the margins of higher education (Ding & Bruce, 2017; Turner, 2011) and adopting a low-status service role (Hyland & Shaw, 2016). Furthermore, the new BLEAP roles in EAP, described by Hadley (2015) are different to those of traditional EAP teachers, one informant in his study describing himself as "teacher, administrative researcher and politician…salesman too" (p. 46).

Although the 'outside view' of EAP provides a contrast to this, it also echoes concerns relating to the marginality of EAP. Many colleagues from the academy appear to consider EAP as being unnecessary (Murray, 2015), of low status (Johns, 1997), a remedial service fix language problems (Turner, 2011) and sub-degree (Fulcher, 2009). However, EAP is also seen as being lucrative and commercial, particularly in the case of summer pre-sessional courses which attract increasing numbers of fee-paying international students (Hyland, 2018), leading to increased vulnerability to outsourcing (Fulcher, 2009). Furthermore, it is unclear as to whether academic colleagues share the view that EAP is an academic discipline and scholarly.

The position of EAP impacts on service delivery and quality in a number of ways. As Bruce (2017, p. 1) suggests, EAP is ‘needs driven’ focusing on the needs of students, researchers and university staff. In order to understand these needs fully, EAP practitioners need to liaise with academic colleagues and become familiar with the specific language conventions and genres of different disciplines. Not having this understanding could lead to many issues, including services which are irrelevant to the real needs of students and staff, curricula of EAP courses which may not match specific disciplinary needs, for example the genres used in assessment, and a lack of support for both home and international students. In addition, a lack of understanding of what EAP can offer may limit the extent to which EAP practitioners are involved in decision-making relating to international students and are able influence the bigger institutional picture (Hyland, 2018). Finally, a misunderstanding of the role of EAP and the potential value to university staff and students may discourage the uptake of services.

The location of EAP may also affect the quality of teaching, due to under-resourcing (Marginson, 2010), tutors with heavy teaching loads (Hadley, 2015) and lack of time or encouragement for scholarship (Ding & Bruce, 2017).

The position of EAP within the academy can be understood through the lens of Pierre Bourdieu's theoretical work on social theory, in particular his concepts of 'field', 'habitus', 'capital' and 'symbolic violence'. Higher education can be analysed as a 'field', or social arena, with its own organisational culture, values, practices and dispositions, or 'habitus'. The hierarchical position of individuals, such as EAP practitioners, within the field of higher education depends on the amount of 'capital' they possess. Bourdieu refers to many forms of capital, including economic capital and social capital (Bourdieu, 1986), but prominent in the

field of higher education are: 'scientific capital' relating to intellectual renown, and 'academic capital' which is related to institutional control (Bourdieu, 1988). Furthermore, power over others is exerted through conscious or unconscious strategies of 'symbolic violence' which soften and disguise domination (Bourdieu, 1992).

An exploration of how EAP practitioners attempt to gain credibility and influence in higher education will be included in this study, focusing on the implications of accruing different forms of capital: scholastic capital in the form of engaging in scholarship (Ding & Bruce, 2017; Swales, 1985), economic capital by focusing on income generation (Fulcher, 2009) or social capital by building social relationships (Field, 2008). Finally, the nature of symbolic violence strategies within higher education will be explored, and their impact on the position of EAP within the academy.

This qualitative study therefore aims to focus specifically on the causes, implications and potential solutions relating to the position of EAP within the academy by addressing a number of key research questions:

- What role does EAP play in marketized HE settings?
- How do EAP practitioners navigate this social space?
- What is the impact of the position of EAP in UK higher education?
- What strategies do EAP practitioners employ to improve their position in the field of higher education?

A qualitative methodological framework has been chosen in order to provide rich data on this complex social issue (Dörnyei, 2007), understand respondent perspectives (Harding, 2019)

and provide a ‘thick description’ (Holliday, 2015) of the situation. A multiple case study research design will be used, and will provide an opportunity to study several EAP units in detail in order to provide an in-depth, holistic and in-context understanding of positioning in organisational structures (Casanave, 2015; Punch, 2016; Yin, 2018).

A sample of three case study sites will be selected representing different organisational positions in higher education. Following a basic mapping exercise to record where EAP services fit within university structures, cases will be selected purposively. It is envisaged that they will be a private provider outside the university structure, an EAP unit based in a faculty within the university, and a unit in a service department, such as an international office. In addition, interviews will be sought with experts in EAP and higher education.

Data will be collected at each site in the form of documentary information and interviews. To facilitate data collection, a data collection protocol will be developed in order to provide an overview of the case study, data collection procedures and research questions (Yin, 2018). Documentary evidence will include documents relating to the university’s international strategy, the EAP unit’s strategy and proposals, records of activity, and web pages relating to international students and language provision. In-depth interviews will be held with key stakeholders at each site, including EAP teachers, EAP managers, academic staff, international office representatives and senior managers responsible for EAP provision. Interviews will be semi-structured with open-ended questions, providing an opportunity to follow-up on areas of interest (Dörnyei, 2007), and will include questions on the positioning of EAP and respondents’ feelings.

Data from documentary evidence and interviews will be analysed using thematic analysis,

taking an inductive approach to identify emerging themes (Harding, 2019). This analysis will use data pooled from all sources and will examine commonalities, differences and relationships in relation to the research questions (Harding, 2019, p. 105). The themes will then be compared to, and located within, Bourdieu's theoretical framework.

A qualitative approach presents issues regarding generalisability of findings and potential researcher bias (Dörnyei, 2007; Hadley, 2017). These will be addressed by selecting diverse research sites and a range of interviewees which represent different experiences of positioning within higher education. Any preconceptions will be analysed early in the project, and documented to provide transparency.

It is hoped that this study will provide a deeper understanding of why EAP services are positioned as they are, how this impacts on quality and service delivery and how EAP can provide more benefit to the institutions where they are placed. This will enable EAP managers and decision makers to make more informed decisions about the location of academic literacy services, and will provide them with strategies for improving their position and impact.

References

Alexander, O. (2012). Exploring teacher beliefs in teaching EAP at low proficiency levels. *Journal of English for Academic Purposes, 11*(2), 99-111. doi:10.1016/j.jeap.2011.12.001

Archer, M. S. (2003). *Structure, agency, and the internal conversation*. Cambridge: Cambridge University Press.

Ball, S. J. (2007). *Education plc: understanding private sector participation in public sector education*. London ; New York: Routledge.

Bell, D. E. (2016). *Practitioners, pedagogies and professionalism in English for Academic Purposes (EAP): the development of a contested field.* (PhD). University of Nottingham.

Benesch, S. (2001). *Critical English for academic purposes: theory, politics, and practice*. Mahwah, N.J. ; London : Lawrence Erlbaum Associates.

Blaj-Ward, L. (2014). *Researching contexts, practices and pedagogies in English for academic purposes*. Basingstoke: Palgrave Macmillan.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education.* (pp. 241-258). New York: Greenwood Press.

Bourdieu, P. (1988). *Homo academicus*. Stanford, Calif. : Stanford University Press.

Bourdieu, P. (1992). *Language and symbolic power*. Cambridge : Polity.

Bruce, I. (2017). *Knowledge and the EAP practitioner: a symposium*. Retrieved from https://teachingeap.files.wordpress.com/2017/01/ian-bruce-baleap-restes-paper-january-28.pdf

Campion, G. C. (2016). 'The learning never ends': Exploring teachers' views on the transition from General English to EAP. *Journal of English for Academic Purposes, 23*, 59-70. doi:10.1016/j.jeap.2016.06.003

Casanave, C. P. (2015). Case studies. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: a practical resources* (pp. 119-135). London: Bloomsbury.

De Vita, G. (2007). Taking stock: an appraisal of the literature on internationalising HE learning. In E. Jones & S. Brown (Eds.), *Internationalising Higher Education* (pp. 154-167). Abingdon: Routledge.

Deem, R., Hillyard, S., & Reed, M. (2007). *Knowledge, higher education, and the new managerialism: the changing management of UK universities*. Oxford: Oxford University Press.

Ding, A. (2016). *Neoliberal EAP: are we all neoliberals now?* Retrieved from https://teachingeap.wordpress.com/2016/01/25/neo-liberal-eap-are-we-all-neoliberals-now/

Ding, A., & Bruce, I. (2017). *The English for academic purposes practitioner*. Basingstoke: Palgrave Macmillan.

Ding, A., & Campion, G. (2016). EAP teacher development. In P. Shaw & K. Hyland (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 547-559). Abingdon: Routledge.

Dörnyei, Z. (2007). *Research methods in applied linguistics : quantitative, qualitative, and mixed methodologies*. Oxford : Oxford University Press.

Ferguson, G. (2007). The global spread of English, scientific communication and ESP: questions of equity, access and domain loss. *Ibérica, 13,* 7-38.

Field, J. E. (2008). *Social capital* (2nd ed.). London : Routledge.

Foskett, N. (2011). Markets, government, funding and the marketisation of UK higher education. In M. Molesworth, R. Scullion, & E. Nixon (Eds.), *The marketisation of higher education and the student as consumer* (pp. 39-52). Abingdon: Routledge.

Fulcher, G. (2009). The commercialisation of language provision at university. In J. C. Alderson (Ed.), *The politics of language education: Individuals and institutions* (pp. 125-146). Bristol: Multilingual Matters.

Hadley, G. (2015). *English for academic purposes in neoliberal universities : a critical grounded theory*. Heidelberg, Germany: Springer.

Hadley, G. (2017). *Grounded theory in applied linguistics research : a practical guide*: Milton Park, Abingdon: Routledge.

Hamp-Lyons, L. (2011). English for Academic Purposes. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 89-105). New York: Routledge.

Harding, J. (2019). *Qualitative data analysis: from start to finish* (2nd ed.). London: Sage Publications.

HESA. (2019). Where do HE students come from? Retrieved from https://www.hesa.ac.uk/data-and-analysis/students/where-from

Holliday, A. (2015). Qualitative research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguisitics: a practical resource* (pp. 49-62). London: Bloomsbury.

Hyland, K. (2018). Sympathy for the devil? A defence of EAP. *51*(3), 383-399. doi:10.1017/S0261444818000101

Hyland, K., & Shaw, P. (2016). *The Routledge Handbook of English for academic purposes*: Milton Park, Abingdon: Routledge.

Jenkins, R. (2008). *Social identity* (3rd ed.). Abingdon: Routledge.

Johns, A. M. (1997). *Text, role, and context : developing academic literacies*. Cambridge: Cambridge University Press.

Marginson, S. (2010). *International student security*. Cambridge: Cambridge University Press.

Moran, M. (2014). *Identity and capitalism*. London: SAGE Publications.

Mowat, J. G. (2015). Towards a new conceptualisation of marginalisation. *European Educational Research Journal, 14*(5), 454-476. doi:10.1177/1474904115589864

Murray, N. (2015). *Standards of English in higher education: issues, challenges and strategies*: Cambridge: Cambridge University Press.

Murray, N. (2016). An academic literacies argument for decentralizing EAP provision. *ELT Journal, 70*(4), 435-443. doi:10.1093/elt/ccw030

Olssen, M., & Peters, M. A. (2005). Neoliberalism, higher education and the knowledge economy: from the free market to knowledge capitalism. *Journal of Education Policy, 20*(3), 313-345. doi:10.1080/02680930500108718

Phillipson, R. (2009). *Linguistic imperialism continued*. London: Routledge.

Punch, K. (2016). *Developing effective research proposals* (3rd ed.): London: SAGE publications.

Raimes, A. (1991). Out of the Woods: Emerging traditions in the teaching of writing. *TESOL Quarterly, 25*(3), 407. doi:10.2307/3586978

Swales, J., M. (1985). ESP - the heart of the matter or the end of the affair? In R. Quirk & E. L. G. Widdlowson (Eds.), *English in the world: teaching and learning the language and literatures.* (pp. 212-223). Cambridge: Cambridge University Press for the British Council.

Turner, J. (2004). Language as Academic Purpose. *Journal of English for Academic Purposes, 3*(2), 95. doi:10.1016/S1475-1585(03)00054-7

Turner, J. (2011). *Language in the academy: cultural reflexivity and intercultural dynamics*. Bristol: Multilingual Matters.

Whitchurch, C. (2013). *Reconstructing identities in higher education: the rise of "third space" professionals*. London: Routledge.

Wingate, U. (2018). Academic literacy across the curriculum: Towards a collaborative instructional approach. *51*(3), 349-364. doi:10.1017/S0261444816000264

Yin, R. K. (2018). *Case study research and applications: design and methods* (6th ed.). Thousand Oaks, California: SAGE Publications.

What can we learn about reading attainment in North East England?

Lindsey Wardle

School of Education, Durham University, Durham, United Kingdom

lindsey.m.wardle@durham.ac.uk

What can we learn about reading attainment in North East England?

This article discusses a forthcoming study which will be focusing on factors which influence reading attainment in North East England. This region is an outlier in UK national assessments; at the end of Key Stage 2, it is the highest attaining region outside of London, but the lowest attaining region nationally at GCSE. Government policy and other organisations have identified literacy on leaving school as a social issue. This study will therefore seek to address the question: What factors help determine the level of reading attainment in North East England? Here, an overview of current evidence relating to the region's attainment is given and gaps in current understanding are identified. This is followed by proposed methodology for the study, which will initially use data from PISA, PIRLS and the National Pupil Database. Finally, the potential impact of this research is discussed, with a focus on how this study will add to the current literature.

Keywords: regional attainment, reading, North East England, PISA

Overview

This article outlines my proposed PhD study which will seek to address the question: What factors help determine the level of reading attainment in North East England? Good literacy skills are associated with higher attainment across other national curriculum subjects and are correlated with higher life expectancy (Nunes et al, 2017; Gilbert et al, 2018), and in the English education system increasing importance is being placed on reading in secondary and particularly primary schools. Attainment in the North East has been a policy issue, with £24million recently invested in the government's Opportunity North East programme, and other initiatives such as the Education Endowment Foundation's 'North East Primary Literacy Campaign' and the National Literacy Trust's 'Read North East' programmes aiming to improve literacy in the region (Gov.uk, 2018; EEF, n.d.; National Literacy Trust, 2017).

The study will conduct a rapid systematic review of evidence, and then analyse linked secondary data from the Programme for International Student Assessment (PISA) and the National Pupil Database (NPD) for regional patterns in reading attainment. In completing a study on this topic, a better understanding of low achievement in a region of high deprivation such as the North East will contribute to knowledge in education and in wider social services. As a conclusion, the study will consider what current policy and provision is doing to improve reading attainment in the North East, whether this is likely to address issues highlighted by the findings, and what might be done in the North East and other areas such as coastal regions and the West Midlands.

Background

As Figure 1 shows, at the end of Key Stage 2, the North East has among the highest percentage of pupils reaching the expected standards in literacy assessments (DfE, 2018b).

On the other hand, at Key Stage 4, the North East clearly has the lowest Attainment 8 and Progress 8 scores for English Language (DfE, 2018a). In short, there is a stark difference in the North East's performance in national assessments at the end of primary school compared to the end of secondary school. Ofsted inspection results reflect this pattern (Watchsted, 2018). According to the National Literacy Trust, only London has more wards in the top three deciles of literacy vulnerability in England than the North East (Gilbert et al, 2018).

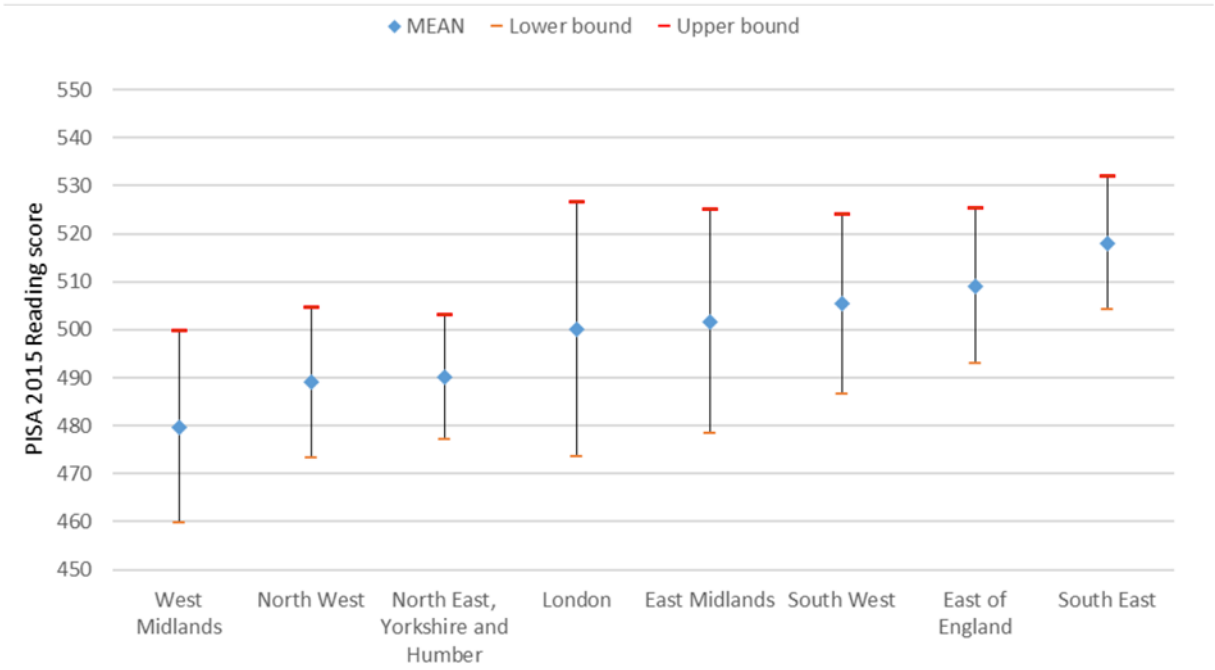
Figure 1: Results by region for pupils at end of Key Stages 2 and 4 (DfE, 2018a; Dfe, 2018b)

Region	% Pupils achieving expected standard in Reading at KS2	% Pupils achieving expected standard in G,P,S at KS2	Attainment 8 Score for English at KS4
Inner London	82	86	10.5
London	81	86	10.6
Outer London	81	86	10.6
North East	81	85	9.5
South East	81	82	10.1
North West	79	83	9.7
South West	79	81	9.9
East	78	80	10.0
West Midlands	77	82	9.7
East Midlands	77	81	9.7
Yorkshire and Humber	76	80	9.6

Internationally, the recurrent Progress in International Reading Literacy Study (PIRLS) presents results for pupils in Year 5, whereas Programme for International Student Assessment (PISA) presents results for pupils aged 15. These international assessments are often cited by government ministers and are used for policy purposes. Since 2006, UK results in these tests have been fairly stable across all subjects, with the UK performing around the OECD average in PISA and above average in PIRLS for reading (NFER, 2018; Jerrim and

Shure, 2016; McGrane et al, 2017; OECD, 2016; IEA, 2017). Despite the importance of these assessments politically, little research has been done using these data at a regional level. For example, Machin et al. (2013) used international data but compared the four UK nations rather than exploring patterns at a regional level. Where these data have been explored at a regional level, in PISA 2015, the North East mean score was among the lowest regionally for reading (Figure 2). Jerrim et al (2017) found a similar pattern to the data found in Figure 1 in mathematics, with high Key Stage 2 and low Key Stage 4 data, and considered this pattern of attainment worthy of further investigation. An equivalent study for reading has not been suggested or completed so far. This new study will fill the gap.

Figure 2: PISA 2015 results by region (Results matched to NPD) (DfE, 2017)



Nationally, reading is becoming an increasingly important focus of policy; reforms in the National Curriculum in 2014 introduced a greater emphasis on phonics and on knowledge, considering that this reading approach will 'raise standards' in schools (DfE, 2015; DfE

2013a; DfE 2013b; DfE 2013c; Marshall, 2017). Moreover, Ofsted has suggested that Key Stage 3 is being overlooked in favour of year groups taking examinations, with inspectors noting examples of slow progress and lack of challenge in English lessons (Ofsted, 2013). A Department for Education report has concluded from national data that progress in Key Stages 2 and 3 is non-linear and non-continuous, but could not explain the reasons for this (DfE, 2011). Hence, there seems to be a mixed picture in national and international assessment data, which warrants further investigation, particularly given the unique pattern found in North East England.

What might the explanations be for differing attainment between regions of the UK at various school stages? Gorard and Smith (2004) argue that there is no evidence that *equivalent* students attain differently in different parts of the UK. This point is further highlighted by Gorard (2018), finding that differences in attainment were predominantly due to pupil characteristics. Alternatively, using data from international assessments, Creese and Isaacs (2016) have identified school-level factors such as centralisation, lack of curriculum integration, and high-stakes tests as flaws in the English system, though the time lag in this data means the full extent of these policy changes are not yet present in data. These are factors that this proposed study will need to consider.

Methods Proposed

To address the factors determining reading attainment in the North East, the following sub-questions will need to be considered:

- What do the major patterns of student attainment in the North East look like?
- How do they compare to other regions of England?

- Are these differences linked with schools types and/or student characteristics? If not, what are the other likely determinants?
- What are the implications for policy and practice in the NE and beyond, and what can be learnt about evidence-led policy-making?

The study will start with a rapid review of the existing literature on the pattern of relative attainment in the NE (high at primary, low at secondary), looking for possible explanations that can be investigated further with the secondary datasets. Search terms will be created and data extracted from included research reports before synthesis. Terms will allude to key themes related to issues of interest linked to the research questions above such as regional differences in outcomes, transition, deprivation, school intakes, reading assessments, SATs, and pupil characteristics. This literature review will inform the secondary data analysis which will follow.

The initial emphasis of this study will be on analysing existing secondary data on reading attainment, with variables for analysis to be determined by the research questions and the review of existing literature detailed above. The scale and availability of secondary data makes it a useful but often under-used resource for researchers, one which can help overcome limitations of other data and where data is already available it is ethical to ensure that this is included in a preliminary investigation (Gorard, 2012; Gorard, 2002). Large-scale secondary datasets can work in tandem with other data to support power analyses (Smith, 2008; Smith, 2012; Gorard, 2002). From the outset, I will need to review what each of the relevant secondary data sources can tell us, including national SATs and GCSE results, the PISA database, and the PIRLS database, beyond the main headline findings aforementioned. I will need to form links between these various sources to form an overall picture of what can be

learned. In bringing together different sources of information in this way, my findings will be stronger, and any discrepancies within this information can be explored later in my study. By doing this, I will be able to discern patterns of student attainment in North East England, and compare these patterns to other regions using the data collated.

The main domain in PISA 2018's survey is reading, so the data released in December 2019 will provide a comprehensive overview during this initial data analysis, with more data items being included related to reading. Using the data from a survey where reading is the main domain will ensure that I can gain the most relevant and up-to-date information on students who will still be in the UK education system. It is essential that data from these sources are understood fully, that their limitations are acknowledged, and that these data are not misinterpreted as has at times been the case (Gorard, 2001). The most recent PIRLS data is from the 2016 survey, and this will also be an important resource; during the course of the study, the PIRLS 2021 data will be released, and I will be able to note any differences in results as I make my conclusions.

Nationally, Key Stage 2 and Key Stage 4 data will continue to be released on an annual basis. Matching PIRLS and PISA data to the NPD will enable me to discover whether the patterns of reading attainment described in the Background section are present in data from independent surveys conducted by the International Association for the Evaluation of Educational Achievement (IEA) and the Organisation for Economic Co-operation and Development (OECD). The timing of this study allows for the most recent international data to be included, and for the National Curriculum changes in 2014 to have been fully implemented in English schools. Assuming there are no further curriculum or assessment changes, this means it is a suitable time to complete a study of this topic. Moreover, current

literature has not used the full extent of data available nor linked international data to the NPD to focus particularly on the North East England region and on reading attainment. All data will be held confidentially, and no individual or organisation will be identifiable.

In order to assess whether school or student differences explain the patterns I find or whether there are alternative explanations, I will need to consider and test to some extent each possible explanation using complex regression modelling. For example, I will need to consider the idea that national assessment data are flawed, particularly in relation to discrimination between performances in Key Stage 2. Measures will need to be considered carefully to avoid identifying the wrong targets and adding unfairness. Regression models will also be suggest any further causal relationships between these independent variables and reading attainment. Any further analysis needed in order to answer the research questions will be decided based on these findings. Ethical considerations will be revisited throughout the study period, and particularly after the initial data analysis when further data will be needed. This will be done with reference to University and to British Educational Research Association's (BERA) Ethical Guidelines (BERA, 2018). At this point, it is important not to bias the study by assuming what the results of the initial data analysis will be; hence, it cannot be stated here what further analysis will be needed beyond that mentioned.

Outcomes

This particular line of research has not been extensively investigated before. Reading is a high priority in current curricula and policy, and the linked data which will be used in this study will be either new or from the most recent secondary datasets available. It will produce good measures of contextual disadvantage based on data from large-scale datasets which will help to identify targets for stakeholders and policy makers. This adds weight to the choice to

research this topic, and should help with trying to disseminate findings. Working with other partners will help with dissemination of findings, something which is already under consideration and being worked towards. Additionally, a successful exploitation of this plan could bring up further questions for research, and provide a model of a suitable research design for further studies on other similar topics such as attainment in science or mathematics.

Acknowledgements

This research is funded by the Economic and Social Research Council through NINE DTP.

References

British Educational Research Association. (2018). *Ethical guidelines for educational research.* (4th ed.). London: BERA. Accessed from: https://www.bera.ac.uk/researchers-resources/publications/ethicalguidelines-for-educational-research-2018

Creese, B. & Isaacs, T. (2016). International instructional systems: how England measures up. *The Curriculum Journal.* 27(1). 151-165.

Department for Education. (2017). *Achievement of 15-year-olds in England: PISA 2015 Further Analysis Report.* Accessed from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/661245/PISA_2015_Further_Analysis_Report_FINAL.pdf

Department for Education. (2018a). *GCSE and equivalent results: 2017 to 2018 (Provisional).* Accessed from: https://www.gov.uk/government/statistics/gcse-and-equivalent-results-2017-to-2018-provisional

Department for Education. (2011). *How do pupils progress during Key Stages 2 and 3?* Department for Education. Accessed from: https://www.gov.uk/government/publications/how-do-pupils-progress-during-key-stages-2-and-3

Department for Education. (2018b). *National curriculum assessments: Key Stage 2, 2018 (Provisional).* Accessed from: https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2018-provisional

Department for Education. (2013a). *National Curriculum in England: English programmes of study – Key Stages 1 and 2.* Department for Education. Accessed from: https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study

Department for Education. (2013b). *National Curriculum in England: English programmes of study – Key Stage 3.* Department for Education. Accessed from: https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study

Department for Education. (2013c). *National Curriculum in England: English programmes of study – Key Stage 4*. Department for Education. Accessed from: https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study

Department for Education. (2015). *Reading: the next steps. Supporting higher standards in schools.* Department for Education. Accessed from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/409409/Reading_the_next_steps.pdf

Education Endowment Foundation. (n.d.). *North East Primary Literacy Campaign*. Accessed from: https://educationendowmentfoundation.org.uk/scaling-up-evidence/campaigns/north-east-literacy-campaign/

Gilbert, L., Teravainen, A., Clark, C. & Shaw, S. (2018). *Literacy and life expectancy: an evidence review exploring the link between literacy and life expectancy in England through health and socioeconomic factors.* National Literacy Trust. Accessed from: https://literacytrust.org.uk/research-services/research-reports/literacy-and-life-expectancy/

Gorard, S. (2001). International comparisons of school effectiveness: the second component of the 'Crisis Account' in England? *Comparative Education.* 37(3). 279-296.

Gorard, S. (2002). The role of secondary data in combining methodological approaches. *Educational Review.* 54(3). 231-237.

Gorard, S. (2012). The increasing availability of official datasets: methods, limitations and opportunities for studies of education. *British Journal of Educational Studies.* 60(1). 77-92.

Gorard, S. (2018). *Education policy: evidence of equity and effectiveness*. Bristol: Policy Press.

Gorard, S. & Siddiqui, N. (2018). There is only research: the liberating impact of just doing research. *International Journal of Multiple Research Approaches*. 10(1). 328-333.

Gorard, S. & Smith, E. (2004). What is 'underachievement' at school? *School Leadership & Management.* 24(2). 205-225.

Gorard, S. & Taylor, C. (2004). *Combining methods in educational and social research.* Maidenhead: Open University Press.

Gov.uk (2018a). Education Secretary launches £24million programme for North East. Accessed from: https://www.gov.uk/government/news/education-secretary-launches-24-million-programme-for-north-east

Gov.uk. (2018b). How to access Department For Education (DfE) data extracts. Gov.uk. Accessed from: https://www.gov.uk/guidance/how-to-access-department-for-education-dfe-data-extracts

IEA. (2017). *What makes a good reader: international findings from PIRLS 2016*. IEA. Accessed from: http://timssandpirls.bc.edu/pirls2016/international-results/pirls/summary/

Jerrim, J., Perera, N. & Sellen, P. (2017). *English education: world class in primary?* London: Education Policy Institute. Accessed from: https://epi.org.uk/publications-and-research/english-education-world-class-primary/

Jerrim, J. & Shure, N. (2016). *Achievement of 15-year-olds in England: PISA 2015 National Report.* Department for Education. Accessed from: https://www.gov.uk/government/publications/pisa-2015-national-report-for-england

Machin, S., McNally, S. & Wyness, G. (2013). Educational attainment across the UK Nations: performance, inequality and evidence. *Educational Research.* 55(2). 139-164

McGrane, J., Stiff, J., Baird, J., Lenkeit, J. & Hopfenbeck, T. (2017). *Progress in International Reading Literacy Study (PIRLS): National Report for England.* Department for Education. Accessed from: https://www.gov.uk/government/publications/pirls-2016-reading-literacy-performance-in-england

National Foundation for Educational Research. (2018). *NFER Education Briefings: Key insights for England from PIRLS, TIMSS and PISA.* NFER. Accessed from: https://www.nfer.ac.uk/nfer-education-briefings-key-insights-for-england-from-pirls-timss-and-pisa/

National Literacy Trust. (2017). Read North East. Accessed from: https://literacytrust.org.uk/communities/north-east/

Nunes, T., Bryant, P., Strand, S., Hillier, J., Barroso, R. & Miller-Friedmann, J. (2017). Review of SES and science learning in formal educational settings. EEF. Accessed from: https://educationendowmentfoundation.org.uk/public/files/Review_of_SES_and_Science_Learning_in_Formal_Educational_Settings.pdf

OECD. (2016). *PISA 2015: Results in focus.* OECD. Accessed from: https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf.

Ofsted. (2015). *Key Stage 3: The wasted years?* Ofsted. Accessed from: https://www.gov.uk/government/publications/key-stage-3-the-wasted-years

Smith, E. (2008). Pitfalls and promises: the use of secondary data analysis in educational research. *British Journal of Educational Studies.* 56:3. 323-339

Smith, E. (2012). Secondary data. In Arthur, J., Waring, M., Coe, R., & Hedges, L. (eds). *Research methods and methodologies in education.* London: SAGE

Watchsted. (2018). Performance tables. Accessed from: https://www.watchsted.com/tables

Task-based Language Teaching in Thai Context: a Call for Robust Evidence

Meechai Wongdaeng

School of Education, Durham University, Durham, the UK

meechai.wongdaeng@durham.ac.uk

Meechai is a doctoral student at the School of Education, Durham University. His doctoral study is funded by the Royal Thai Government. He has worked in a public university in the south of Thailand for several years. His research interest is to apply evidence-based research to inform educational policy and practices.

Task-based Language Teaching in Thai Context: a Call for Robust Evidence

The task-based language teaching (TBLT) is a learner-cantered pedagogical approach which promotes learners' engagement in communicative tasks. TBLT has been implemented widely including in Thailand where English is used as a foreign language. However, the evidence of its effectiveness remains equivocal. This review was conducted to examine and synthesize the evidence of TBLT benefits in EFL contexts. However, it was found that most TBLT research in EFL contexts aiming to draw the effects of TBLT intervention tends to be predominated by low rigorous designs. For Thailand in particular, most studies report students' positive perceptions about the TBLT interventions while the claims to improve language competences are vaguely proved due to the low evidence-based rigor. From such review-based findings, the paper proposes a call for design-based research to evaluate the impact of TBLT on language competences and learning skills. Adopting rigorous designs which provide counterfactuals would produce a more secure evidence for policy and practices of the TBLT implementation in the Thai EFL context.

Keywords: Task-based language teaching, systematic review, English as a foreign language, evidence-based evaluation

1. Introduction

The task-based learning has been widely applied in many domains including English language education as an approach to promote more active role of students in the learning process (Nunan, 2004). The approach has also been implemented by throngs of EFL teachers and researchers with an aim to improve English proficiency of the learners. Despite its promises and popularity, the effectiveness of task-based language teaching (TBLT) in EFL contexts is still not definitive and there are concerns over the cultural barriers which might impede the implementation of TBLT in EFL contexts (e.g. Butler, 2011; McDonough and Chaikitmongkol, 2007). Therefore, there is a need for clearer evidence on the effectiveness of TBLT for EFL learners.

Focusing on the Thai context, English proficiency of Thai citizens is considered one of the key factors to keep the country competitive in the global economy, leading to a huge investment by the government on English education (Hayes, 2016). The official establishment of the Association of South East Asian Nations (ASEAN) as one economic community in 2015, made English, which is declared as an official lingua franca of the region, even more important to Thais especially the young and the working-age generations.

Numerous studies have been conducted to improve the outcome of English education in Thailand and TBLT has been one of the approaches for this attempt. Nonetheless, the effectiveness of TBLT in Thai context has hardly been robustly assessed. Therefore, the main aim of this paper is to investigate the evidence of TBLT in EFL and Thai contexts.

2. The Fundamentals of TBL in English Language Teaching

2.1 Background to TBLT

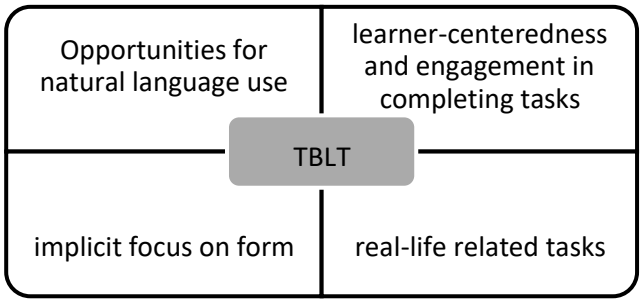
Since its emergence in the mid-1970s, TBLT has played an influential role in language education and has been an established area of research (Samuda, Bygate & Van den Branden, 2018). Central to TBLT, a clear understanding of the term 'task' is important. Tasks are defined broadly by some scholars as language activities which engage the learners in using the target language (e.g. Prabhu, 1987; Littlewood 2004). Others argue that engagement is necessary but insufficient in TBLT. Willis (1996 p.23) defines tasks as 'activities where target language is used by the learners for a communicative purpose in order to achieve an outcome' (p.23). By this definition, the task is essentially a communicative activity with an aim to achieve the outcome. Ellis (2003) maintained that tasks are not simply any activities but need to be related to authentic purposes in communicative situations so that they can be a workplan which necessitates language use as a mean for task accomplishment. Nunan (2004) divided 'tasks' into *target tasks* and *pedagogical tasks* and it is the latter which is focused in the classroom to "involve learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning" (p.4).

The multiple definitions related to tasks represent the two orientations of TBLT. The 'task-based' is regarded as a strong form of TBLT which puts task as the core of syllabus design while the 'task-supported' is viewed as a weak form of TBLT which supports learning of language functions through communicative language tasks (Samuda, Bygate & Van den Branden, 2018). Indeed, both orientations of TBLT have a role to play in the real classroom practice, rather than one being superior to the other. The strong orientation has been prevalent in the English as a second language (ESL) contexts (Thomas, 2017). However, for EFL contexts where there are little opportunities for English use outside the classrooms, the adaptation of TBLT or its weak form can be appropriate (Carless, 2009; Butler, 2011).

2.2 Principles of Task-Based Language Teaching

The key conceptual basis of the TBLT is based on the theory of experiential learning by Dewey (1938) which emphasizes learner's participation and collaboration in the community of leaning in order to solve real-world problems (Norris, 2009). From this concept, two rationales can be inferred about TBLT. First, language is not only learned *in order* to use it functionally but it is learned *by* making functional use of it. Second, there should be a close link between the task which students do and the real-life situations outside the classroom (Van den Branden, 2006 p.6).

Ellis (2009) suggests that there is no single way to implement TBLT. However, he proposed four characteristics and one optional feature of TBLT as follows;



(Ellis, 2009 p.225)

Another characteristic which is optional in TBLT is the traditional structural teaching which can be rejected or can be included to complement the communicative TBLT.

On the principles of TBLT, Nunan (2004) proposed that the task-based approach should be based on seven principles (p.35);

Scaffolding: providing a supporting framework for the learners to an appropriate extent

Task dependency: sequence of task elements relates and builds on each other

Recycling of language: reintroducing the target content over a period of time

Active learning: learners learn best when actually involving in or using the target language

Integration: Showing to the learners the relationship between language form, communicative function and meaning of the target content

Creative reproduction: learners produce the language model in novel ways

Reflection: learners reflect on what they have learned and how well they are doing.

From the principles and characteristics offered by Ellis (2009) and Nunan (2004), it could be noted that the opportunities for natural language use, engagement in real-life tasks, scaffolding and reflection are important considerations in the TBLT. To accommodate these principles, Norris (2009 p.583) proposed a four-phase procedure of task-based instruction detailed as follows;

1. **Task input:** Introduce the target task as it is actually used in the real-world situations.
2. **Pedagogic task work:** Tasks are elaborated and manipulated to raise learners' awareness of new language forms and functions.
3. **Target task performance:** Encourage the learners to accomplish the target tasks in communicative situations.
4. **Task Follow-Up:** Teachers and learners reflect on the performance in the previous phase in terms of language, content, task knowledge.

This procedure was chosen to discuss here as it allows the learners to engage in real life tasks and learn to communicate through task accomplishment as it should be in TBLT. Moreover, it scaffolds the learners from less demanding tasks to a more demanding one. In addition, it recognizes the importance of the reflective thinking emphasized in the task follow-up. Thus, it is an appropriate approach to transfer all key TBLT principles for developing learners' competence. It can better promote learners' engagement and thinking than the traditional three-stage procedure of pre-during-post tasks.

3. English Language Teaching in Thailand

English language policy in Thailand has gone through multiple changes. Since the education reform through the National Education Act 1999, English curriculum seemed to adopt more communicative approach and promote more learner-centred environment (Wongsothorn et. al., 2002). However, the classroom teaching was still content-based and form-focused instruction still seemed to be the norm. The curriculum was revised again in 2008 to provide clearer goals and standards and allow more freedom for teachers in syllabus design and pedagogic methodology (Nonthaisong, 2015). Unfortunately, the move has not managed to enhance the outcomes of English learners in Thailand (Baker and Jarunthawatchai, 2017).

A large amount of research has been undertaken to improve the standards of English language education in Thailand and a myriad of TBLT studies have been reported (e.g. McDonough & Chaikitmongkol, 2007; Wongdaeng & Hajihama, 2018). However, the TBLT has still been unable to make an observable impact on English education. One explanation for the poor performance can be due to the quality of education research and

accountability system (UNESCO Bangkok, 2017). This reflects the need for higher quality of research on English education in the Thai context.

4. Rationales for Undertaking This Review

The research on TBLT has attracted a lot of EFL researchers and has been carried out with different focus and age groups of learners. Most of the findings agree that students have positive attitudes towards TBLT. However, the impact of TBLT on skills improvement still seem unclear. To evidently find out the effectiveness of this approach, it is important that robust research designs which can provide counterfactual information are used (Shadish, Cook & Campbell, 2002). This provides a rationale for this paper to conduct a systematic review of controlled trials implementing TBLT interventions in EFL contexts and in Thailand to assess the evidence of the TBLT effectiveness. The explicit, transparent, replicable method in the systematic reviews can minimise biased views in the findings (Torgerson, Hall & Light, 2012). The literature reviews in ELT research need to be based on a more explicit approach rather than on an arbitrary selection of the studies to be reviewed (Low and Beverton, 2004).

From the contextual ground, systematic reviews of TBLT interventions in Thai contexts have never been undertaken despite numerous reports of its implementations. A review which applies a systematic approach can shed more light on the TBLT effectiveness. The promotion of ELT policy and practices in Thailand needs to be more evidence-based instead of being motivated by subjective preferences or ad hoc political agenda.

5. Design and Methods

A systematic approach was used for reviewing the relevant literature on the TBLT interventions in EFL contexts in order to investigate the potential effectiveness of TBLT among the EFL learners. The review addresses the following questions;

1. What is the evidence of the effectiveness of TBLT on language competence or study skills of the English learners in EFL contexts?
2. What is the evidence of the effectiveness of TBLT on language competence or study skills of Thai EFL learners?

5.1 Systematic search

The ScienceDirect, the Education Resources Information Centre (ERIC) and the Thai Journal Online (ThaiJo) databases were selected to search for the relevant studies because the ScienceDirect and ERIC are the two databases which EFL researchers commonly publish their works and the ThaiJo is the biggest online research database in Thailand. The publication date range was limited to 2004 to 2018. The search was undertaken between 15-20 May 2019.

Table 1 Search Strategies

Databases	Boolean string	Number of hits
Science Direct	("task-based" OR "Project-based") AND ("English as a foreign language" OR “EFL”) AND ("experimental" OR "effect")	342
Education Resources Information Center (ERIC)	"task-based" OR "Project-based" + "English as a foreign language" OR “EFL”+ "experimental" OR "Effect"	54
Thai Journal Online (ThaiJo)	- task-based + English -project-based + English	32 19
TOTAL		447

Note: The term ‘randomised controlled trial’ was not used in the search to allow as many search results as possible.

5.2 Identification of studies

After the search, the studies were identified by screening titles and abstracts to determine whether they met the PICO eligibility criteria in Table 2.

Table 2 Inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
a. Population	- learners of English as a foreign language	-learners of English in the in the English-speaking or ESL settings
b. Intervention	-focus on pedagogy design of Task-based language teaching or Project-based language teaching with/without use of technology	-Use of technology-enhanced task but focus on the use of technology instead of the pedagogical approach
c. Comparison	-include a comparison group	-pre/post, non-comparison group
d. Outcomes	-language competences or study skills	-perceptions, satisfactions, motivation

After the first screening, 18 studies from ScienceDirect, 18 from ERIC and 2 from ThaiJo remained eligible. In the second screening, the remaining studies were skimmed through the whole text based on the PICO criteria. Six more studies were excluded at this stage, leaving 32 studies for the data extraction (see Figure 1).

It should be noted that the screening was conducted by the researcher. Indeed, having one or more researchers collaborating as an inter-rater in the screening of studies would help make the selection more reliable.

5.3 Data Extraction and quality appraisal

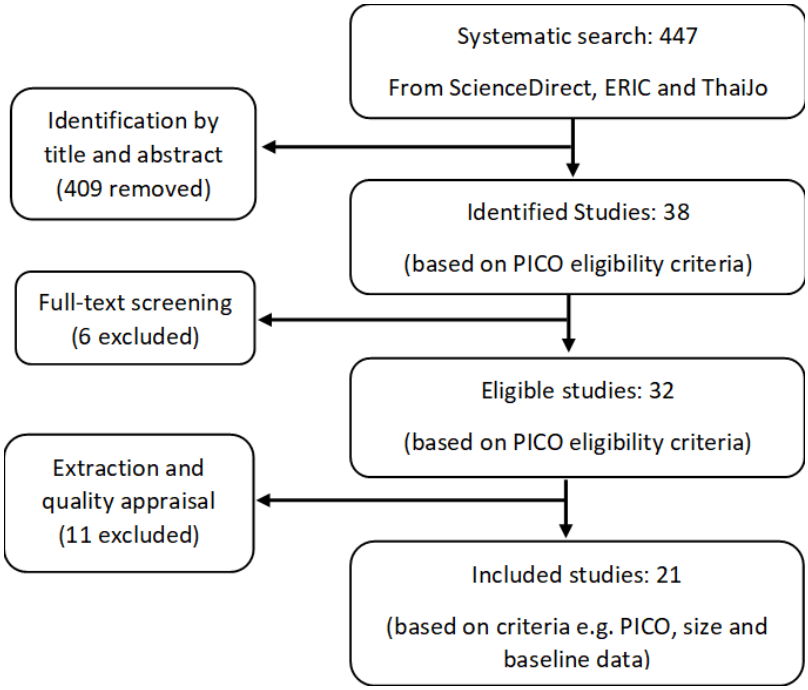
The remaining studies were fully scanned to extract data on the setting, intervention, controlled group condition, outcome variables, and outcomes measures. At this stage, the remaining studies were quality appraised based on the PICO criteria with additional quality criteria in terms of sample size and baseline data. These two issues were used to further screen the studies because the sample size can affect the findings and weak baseline data can undermine the comparability of the intervention groups. 11 more studies were excluded, leaving 21 studies included for the data synthesis. Summary of the identified studies are presented in Appendix 1.

6. Results

6.1 Search results

The results from each of the review processes are presented in figure 1.

Figure 1. Flow diagram of the review process



6.2 Synthesis of the findings

This section will present a synthesized finding from the included studies. There are a number of options for synthesising the findings in a systematic review such as narrative synthesis, vote-counting and meta-analysis. To choose an appropriate approach for the synthesis, factors such as review questions, and the homogeneity of the studies should be considered (Torgerson, 2003). In the vote-counting approach, the findings of the studies are identified whether they report the positive statistical significance or negative or neutral and are counted to get the most typical results which represent the overall effect of the intervention (Cook et. al., 1992). Despite some criticisms of failure to recognise the different characteristics of the included studies and their differing methodological rigor, vote-counting can be useful for describing the overall effect of the relevant studies especially when a meta-analysis is not

possible (Davies, 2000). The vote-counting and narrative approaches will be used in this review because the outcome measures in the included studies focus on different language skills which require heterogenous assessment methods such as the objective tests in reading and grammar and the criteria-based judgement in writing tests. The sample size and methodological rigor of the included studies will be considered in the discussion.

6.2.1 Effects of TBLT on EFL learners' language achievement

Table 3 Synthesis of the effects of TBLT interventions on EFL learners' language achievement

Study reference	**Outcome**	**Sample size** (Int./Cont.)	**Post-test mean** (Int./Cont.)	**SD** (Int./Cont.)	**Effect size**	**Supports TBLT**
Kafipour et al. (2018)	Written production	40/40	Content 2.486/1.794 Organization 1.929/1.632 Vocabulary 2.627/2.000 Language 2.771/2.441 Mechanics 2.829/2.353	Content .7724/.6169 Organization .5021/.5267 Vocabulary .4902/.4083 Language .4260/.4527 Mechanics .4363/.5154	-	yes
Chou (2017)	Listening	44/44	71.07 / 66.07	7.93 / 11.15	0.5168	yes
Madhkhan & Mousavi (2017)	Reading comprehension	70 (unclear allocation)	n/a	n/a	-	yes
NamazianDost et al. (2017)	Grammatical achievement	40/40	33.80 /28.60	2.45158 / 2.22803	2.2199	yes
Azizifar et al. (2015)	Reading comprehension	30/30	15.78 / 13.25	4.57 / 4.66	0.5482	yes
Amirian & Abbasi (2014)	Grammar competence	31/31	14.4194/11.6774	3.74855 / 2.91418	0.8167	yes
Marzban & Hashemi (2013)	Speaking	32/32	71.75 / 75.28	8.056 / 6.517	-0.482	no
Setayesh & Marzban (2017)	Reading comprehension	Int.1.=25 Int.2=25 Cont.2=25 Cont.3=25	32.54 36.59 25.75 26.73	7.43 7.86 8.76 6.45	-	yes
Shiraz & Larsari (2014)	Reading comprehension	Int.1=40 Int.2=40 Cont.=40	36.32 36.90 32.15	3.682 2.808 3.042	-	yes
Tilfarlioglu & Basaran (2007)	Reading comprehension	28 / 28	n/a	n/a	-	yes

Study reference	Outcome	Sample size (Int./Cont.)	Post-test mean (Int./Cont.)	SD (Int./Cont.)	Effect size	Supports TBLT
Shabani & Ghasemi (2014)	Reading comprehension	30/30	67.16 59.80	8.27 10.85	0.763	yes
Khodabandeh (2016)	Writing classified ads	Self=18 / Exp=18 Imp=18 Task=18	n/a	n/a	-	yes
Saeheng & Prammanee (2012)	Reading comprehension	20/20	76.65 / -	-	Use 80% criteria	yes

Note: n/a = the studies provide other statistics, not reporting mean and S.D

Table 3 presents the findings from studies which investigated the effect of TBLT intervention on language learning achievement. From the thirteen relevant studies, most of them report positive effect of TBL on different language skills. One study by Marzban & Hashemi (2013) report non-significant effect of the TBL intervention. Overall, the TBLT interventions have positive effects for improving language learning among EFL learners.

6.2.2 Effects of TBLT on specific language skills

Considering specific language skills, the benefit of TBL interventions are mostly explored in the reading and writing skills. Speaking, listening and vocabulary skills are still lacking empirical interventions, indicating a need for more robust TBLT studies on these language areas.

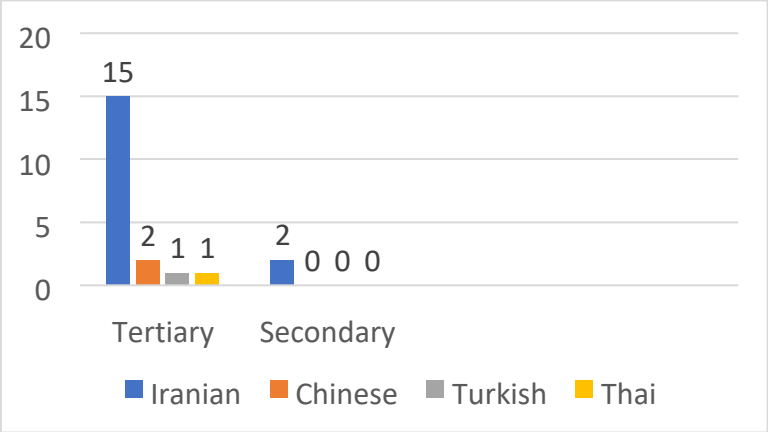
Table 4 Effects of TBLT on specific language skills

Skills	Reading	Writing	Grammar	Listening	Vocab	Speaking
Number of studies	8	5	3	3	2	1
Effect of TBLT	positive	positive	positive	positive	positive	negative

6.2.3 Research contexts

Most studies were conducted with EFL learners in tertiary education and 2 studies recruited secondary school students. Most studies were in Iranian contexts while only one Thai study met the eligibility criteria (see Figure 2). The excluded studies were those with weak design due to lacking a comparison group. Without counterfactual evidence from the comparator, the claimed impact of the intervention is unwarranted (Gorard, 2013).

Figure 2. Contexts of the included studies



6.2.4 Other variables found in the included studies

Task type and complexity

-The high complexity of task has no impact on the target language skills. Thus, simple tasks are preferable (Attarzadea & Farahani, 2014; Kasiria & Fazilatfarb, 2016).

-Tasks which allow strategic planning are better than tasks with no plan (Asgarikia, 2014) and collaborative tasks are better than individual tasks (Zareia & Naamaeib, 2014).

Integration of technology

-Task-based instruction is better than simply teaching using technology (Rajabia & Hashemiana, 2015). When technology is integrated in TBLT, it is effective in improving the target skills (Mohamadi, 2018; Tian & Suppasetseree, 2013).

Metacognition

The task-based instruction can improve the learners' language skills as well as their metacognitive awareness especially in planning and evaluation, directed attention and problem-solving (Chou,2017).

6.3 Quality appraisal

Despite all the included studies meeting the inclusion criteria, there are a few issues which undermine the strength of the synthesized evidence. Firstly, because the participants in the studies were not allocated to groups randomly, the small sample size in many studies (i.e., Rajabi & Hashemiana, 2015; Saeheng & Prammanee, 2012; Tilfarlioglu & Basaran, 2007) can exaggerate the results (Coe, 2002). In terms of quality, some studies do not provide clear information about the participants and group allocation (i.e., Attarzade & Farahani, 2014; Kasiri & Fazilatfar, 2016; Madhkhan & Mousavi, 2017). This can minimise the trustworthiness of the findings as the information is important for validity check (Torgerson, 2003). In reporting findings, some studies do not consistently provide the results from the

controlled groups, making the counterfactual information weak (i.e., Saeheng & Prammanee, 2012).

7. Discussion

The synthesized findings in table 3 seem to show the effectiveness of TBLT in EFL context. However, with quality issues discussed earlier, the task-based approach is highly promising for improving learning achievement of EFL learners but the evidence of the effect is still inconclusive. Moreover, the findings are strongly dominated by the studies in Iranian contexts. This emphasizes the need for more rigorous studies in Thai EFL contexts to provide clearer evidence on the topic.

The findings from the review may not collocate with McDonough and Chaikitmongkol (2007)'s study which reported difficulties experienced by Thai learners in task-based setting. The task-based approach has been through a lot of trial-and-error implementations in EFL contexts as one of the ways to improve English proficiency of the students (Kettanun, 2015). With its widespread application, it is appropriate to rigorously evaluate the impact of TBLT in order to provide evidence for policy and pedagogical practices. With the more influential roles and emphasis of English language in Thailand after the integration of ASEAN community in 2015, it is necessary for the research on TBLT and other interventions to be more rigorously conducted and evaluated if the TBLT research is to improve or inform about English language education in this country.

8. A Way Forward

This section proposes a few suggestions about the future research on TBLT effectiveness in Thai context. Firstly, the review of previous TBLT studies suggests that the impact evaluations of TBLT implementation in Thailand are critically lacking. Despite numerous studies aiming to investigate the effect of TBLT, only one TBLT study was found to meet the PICO criteria, leaving the non-comparator studies excluded. Strong research designs which provide counterfactual evidence are essential for establishing a link between the intervention and the effect (Gorard, 2013). The evidence-based research could provide more trustworthy findings which could be used to inform policy and pedagogical practices (HM Treasury, 2011).

Secondly, the evidence of TBLT on listening comprehension is insufficiently assessed. The listening skill is fundamental to understanding and communication but is often overlooked, compared to other language skills (Goh, 2008) and this is also true in ELT research in Thailand (Woottipong, 2014). Thus, TBLT research which targets listening skills would fill the gap in the EFL literature and satisfy the need of oracy development among Thai EFL learners.

Other useful aspects for the TBLT research are the appropriate incorporation of technology and the development of metacognitive awareness. The ability to use the ubiquitously available technology is a key 21st century skill which the learners should be capable of and so should the teacher (González-Lloret & Ortega, 2014). The studies included in this review

(e.g. Mohamadi, 2018) suggested that applying technology in TBLT is an effective way of instruction. This is because the very characteristics of TBLT such as peer interaction, collaboration and experiential learning lend itself for the incorporation of technological application (González-Lloret & Ortega, 2014). As Higgins et. al. (2012) suggested, merely using technology in teaching doesn't guarantee effectiveness and it needs to be applied with an appropriate pedagogy.

The metacognitive instruction was found to improve English listening and metacognitive awareness in Chou (2017). It was the only study among the included trials which considers metacognition. However, it is of relatively high quality, compared to other included studies in such areas as group size, design and report of the relevant information. Moreover, metacognitive instruction has been widely reported in several syntheses to be effective for improving learning (e.g. EEF, 2018; Higgins et.al., 2005; Plonsky, 2011). The metacognitive task-based interventions can be a promising approach for improving English education in Thailand and rigorous studies are needed to evaluate the impact of the approach.

9. Conclusion

The review has found the promising impact of TBLT for learners in EFL contexts as most studies report positive effect of TBLT on students' achievement. However, the evidence deems equivocal due to the limited evidence-based rigor of the included studies. This emphasizes the necessity for TBLT research in EFL contexts to be conducted more robustly paying more careful attentions to the internal validity. The listening skill is minimally explored and requires more investigations. Along with language proficiency, considerations on technology integration and metacognitive development can be intriguing variables to be assessed.

References

Baker, W., and Jarunthawatchai, W. (2017). English language policy in Thailand. *European Journal of Language Policy, 9*(1), 27–44.

Butler, Y. G. (2011). Implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics, 31*, 36–57.

Carless, D. (2009). Revisiting the TBLT versus PPP debate: voices from Hong Kong. *Asian Journal of English Language Teaching, 19*, 49-66.

Cook, T.D., Cooper, H., Cordray, D.S., Hartmann, H., Light, R.J., Louis, T.A. & Mosteller, F. (1992). *Meta-Analysis for explanation*. New York: Russell Sage Foundation.

Davies, P. (2000). The relevance of systematic reviews to educational policy and practice. *Oxford Review of Education, 26* (3-4), 365-378.

Dewey, J. (1938). *Experience and education*. Now York: Collier Books.

Education Endowment Foundation (EEF). (2018). Teaching and learning toolkit. Retrieved July 15, 2019, from https://educationendowmentfoundation.org.uk/evidence summaries/teaching-learning-toolkit/

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.

Ellis, R. (2009). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221-246.

Goh, C. (2008). Metacognitive instruction for second language listening development: theory, practice and research implications. *RELC Journal, 39*(2), 188-213.

González-Lloret, M. & Ortega, L. (2014). Towards technology-mediated TBLT. In M, González-Lloret & L, Ortega (eds.*), Technology-mediated TBLT: researching technology and tasks* (pp. 1-22). Philadelphia: John Benjamins.

Gorard, S. (2013). *Research design: creating robust approaches for the social sciences*. London: Sage.

Hayes, D. (2016). The value of learning English in Thailand and its impact on Thai: perspectives from university students. *Asia Pacific Journal of Education, 36*(1), 73-91.

Higgins S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A Meta-analysis of the impact of the implementation of thinking skills approaches on pupils*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Higgins, S., Xiao, Z. & Katsipataki, M. (2012). *The Impact of digital technology on learning: a summary for the Education Endowment Foundation*. London: Education Endowment Foundation.

HM Treasury. (2011). *The Magenta Book: guidance for evaluation*. London: HM Treasury.

Kettanun, C. (2015). Project-based learning and its validity in a Thai EFL classroom. *Procedia - Social and Behavioral Sciences, 192,* 567 – 573.

Littlewood, W. (2004). The task-based approaches: some questions and suggestions. *English Language Teaching Journal. 52*(4), 319-326.

Low, G., and Beverton, S.(2004). *A systematic review of the impact of ICT on literacy learning in English of learners between 5 and 16, for whom English is a second or additional language*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

McDonough, K., and W. Chaikitmongkol (2007). Teachers' and learners' reactions to a task based EFL course in Thailand. *TESOL Quarterly 41,* 1071–1132.

Nonthaisong, K. (2015). *The impact of English language policy in the Basic Education Core Curriculum B.E. 2551 (A.D. 2008) on English instruction of Thai secondary EFL teachers: a multiple case study in Northeast Thailand*. (doctoral dissertation). The University of Alabama, Alabama, USA.

Norris, J. M. (2009). Tasked-based teaching and testing. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 578–594). Malden, MA: Wiley-Blackwell.

Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: a meta-analysis. *Language learning, 61*(4), 993-1038.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Basingstoke: Palgrave Macmillan.

Samuda V., Bygate M., & Van den Branden, K. (2018). *TBLT as a researched pedagogy*. Amsterdam: John Benjamins.

Shadish, W.R., Cook T. D. and Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Torgerson, C. (2003). *Systematic reviews*. London: Continuum.

Torgerson, C., Hall, J. & Light, K. (2012). Systematic reviews. In J. Arthur, M. Waring, R. Coe & L.V. Hedges (eds), *Research methods and methodologies in education* (pp.217-230). London: SAGE.

Thomas, M. (2017). *Project-based language learning with technology: learner collaboration in an EFL classroom in Japan*. NY: Routledge.

UNESCO Bangkok. (2017). *2017/2018 Global education monitoring report: Thailand highlights*. Retrieved Aug 5, 2019, from https://bangkok.unesco.org/content /20172018-global-education-monitoring-report-thailand-highlights

Van den Braden, K. (2006). Introduction: task-based language teaching in a nutshell. In K.Van den Braden (ed), *Task-based language education: from theory to practice (pp. 1- 16)*. Cambridge: Cambridge University Press.

Willis, J. (1996). *A framework for task-based learning*. Harlow, UK: Longman.

Wongdaeng, M and Hajihama, S. (2018). Perceptions of project-based learning on promoting 21st Century skills and learning motivation in a Thai EFL setting. *Journal of Studies in the English Language, 13* (2), 158-190.

Wongsothorn, A., Hiranburana, K. and Chinnawongs, S. (2002). English language teaching in Thailand today. *Asia Pacific Journal of Education, 22*(2,) 107-116.

Woottipong, K. (2014). Effect of using video materials in the teaching of listening skills for university students. *International Journal of Linguistics, 6*(4), 200-212.

References of Reviewed Studies

Amirian, S. and Abbasi, S. (2014). The effect of grammatical consciousness-raising tasks on Iranian EFL learners' knowledge of grammar. *Procedia - Social and Behavioral Sciences, 98*, 251-257.

Asgarikia, P. (2014). The effects of task type, strategic planning and no planning on written performance of Iranian intermediate EFL learners. *Procedia - Social and Behavioral Sciences, 98*, 276-285.

Attarzade, S. and Farahani, E. (2014). The effect of task complexity on Iranian EFL learners' listening comprehension across aptitude. *Procedia - Social and Behavioral Sciences, 98*, 314-323.

Azizifar, A., Babaei, M., Jamalinesari, A. and Gowhary, H. (2015). The effect of grammatical consciousness raising task on Iranian EFL learners' reading comprehension. *Procedia - Social and Behavioral Sciences, 192*, 252-259.

Chou, M. (2017). A task-based language teaching approach to developing metacognitive strategies for listening comprehension. *International Journal of Listening, 31*(1), 51- 70.

Kafipour, R., Mahmoudi, E. and Khojasteh, L. (2018). The effect of task-based language teaching on analytic writing in EFL classrooms. *Cogent Education, 5*(1),1-16.

Kasiri, F. and Fazilatfar, A. (2016). The impact of task complexity on cognitive processes of L2 writers and writing quality: the case of writing expertise, L1, and lexical retrieval. *Procedia - Social and Behavioral Sciences, 232*, 561-568.

Khodabandeh, F. (2016). Comparing the effects of four instructional treatments on EFL students' achievement in writing classified ads. *English Language Teaching, 9*(3), 139-152.

Madhkhan & Mousavi (2017). The Effect of Implimentation of TBLT in Reading Comprehension Classes of Iranian EFL Learners. *English Language Teaching, 10* (11), 119-128.

Marzban, A. and Hashemi, M. (2013). The impact of opinion-gap tasks on the speaking of Iranian intermediate EFL learners. *Procedia - Social and Behavioral Sciences, 70*, 943-948.

Marzban, A. and Mokhberi, M. (2012). The effect of focus on form instruction on intermediate EFL learners' grammar learning in task-based language teaching. *Procedia - Social and Behavioral Sciences, 46*, 5340-5344.

Mohamadi, Z. (2018). Comparative effect of project-based learning and electronic project-based learning on the development and sustained development of English idiom knowledge. *Journal of Computing in Higher Education, 30*(2), 363-385.

NamazianDost, Bohloulzadeh & Pazhakh (2017). The effect of task based language teaching on motivation and grammatical achievement of EFL Junior High School students. *Advances in Language and Literary Studies, 8* (2), 243-259.

Rajabi, B. and Hashemian, M. (2015). TBLT or technology: which is effective to teach resumptive pronouns in writing?. *Procedia - Social and Behavioral Sciences, 192*, 321-326.

Saeheng, P. and Prammanee, N. (2012). The effect of using a task-based interactive learning program on English reading ability of Higher Vocational Certificate accounting students. *HRD Journal, 3* (1) 63-70.

Setayesh, M. and Marzban, A. (2017). The impact of task-based language teaching on the development of Iranian EFL learners' ESP reading comprehension skills. *Advances in Language and Literary Studies, 8*(2),70-76.

Shabani, M. and Ghasemi, A. (2014). The effect of task-based language teaching (TBLT) and content-based language teaching (CBLT) on the Iranian intermediate ESP learners' reading comprehension. *Procedia - Social and Behavioral Sciences, 98*, 1713-1721.

Shiraz & Larsari (2014). The effect of project-based activities on intermediate EFL students' reading comprehension ability. *The Journal of Effective Teaching, 14* (3), 38-54.

Tian, X. & Suppasetseree, S. (2013). Development of an instructional model for online task-based interactive listening for EFL learners. *English Language Teaching, 6* (3), 30-41.

Tilfarlioglu, F. & Basaran, S. (2007). Enhancing reading comprehension through task-based writing activities: an experimental study. *The Reading Matrix, 7*(3), 134152.

Zarei, A. and Naamaei, F. (2014). The effect of pre-task activity types (SRE, PALS, and CSR) on L2 reading comprehension and vocabulary recognition and recall. *Procedia - Social and Behavioral Sciences, 98*, 1996-2003.

Appendix A: Characteristics of included studies

Study reference	**Settings**	**Interventions**	**Control**	**Outcome variable**	**Sample size** (E/C)
Kafipour et al. (2018)	Intermediate Iranian Tertiary	Task-based Wring Instruction	Business as usual	Writing	40/40
Chou (2017)	Intermediate Chinese Tertiary	Strategy-embedded Task-based framework for listening	strategy-based instruction	Listening and metacognition	44/44
Madhkhan & Mousavi (2017)	Iranian Tertiary	20 sessions of Taskbased reading instruction	Business as usual	Reading	70 (unclear allocation)
NamazianDost et al. (2017)	Iranian Tertiary	12 sessions of TBLT	Business as usual	Grammar	40/ 40
Azizifar et al. (2015)	Secondary School Iranians	six-week Grammatical Consciousness Raising task	Business as usual	Reading comprehension	30/30
Amirian & Abbasi (2014)	Secondary School Iranians	Grammar Consciousness Raising Task	PPP method	Grammar	31/31
Marzban & Hashemi (2013)	Adults Iranians	10-week Opinion gap task-based instruction	Business as usual	speaking	32/32
Setayesh & Marzban (2017)	Iranian EAP tertiary	4-week TBLT instruction	Grammar translation Method	Reading comprehension	Int.1.=25 Int.2=25 Cont.2=25 Cont.3=25
Shiraz & Larsari (2014)	Intermediate Iranian Tertiary	2 groups of PBL instruction	Communicative language teaching (CLT)	Reading comprehension	Int.1=40 Int.2=40 Cont.=40
Tilfarlioglu & Basaran (2007)	Turkish Tertiary	TBLT	Business as usual	Reading comprehension	28/28
Shabani & Ghasemi (2014)	Iranian Tertiary	11 sessions of TBLT	CBLT	Reading comprehension	30/30
Khodabandeh (2016)	Iranian Tertiary	Self-study treatment	3 comparison groups: -explicit teaching -implicit teaching - task-based instruction	Writing classified ads	Self=18 / Exp=18 Imp=18 Task=18

Saeheng & Prammanee (2012)	Thai Higher Vocational Students	18 sessions of Taskbased instruction	business as usual	Reading comprehension	20/20
Attarzade & Farahani (2014)	Upper intermediate Iranian Tertiary	Task-based with higher task complexity	Task-based with lower task complexity	Listening	70 Low 58 High 12
Kasiri & Fazilatfar(2016)	Iranian Tertiary	Task-based with higher task complexity	Task-based with lower task complexity	Writing	60 (unclear group allocation)
Asgarikia (2014)	Iranian Tertiary	Narrative Writing Tasks strategic planning	Narrative writing task with no strategic planning	Writing	30/30
Marzban & Mokhberi (2012)	Adults Iranians	Two groups • reactive Focus on Form Task • pre-emptive Focus on Form Task	Business as usual	Grammar learning	P27 R26 C26
Zarei & Naamaei (2014)	Iranian Tertiary	Three task types 1 Scaffolded Reading Experience 2.Collaborative Strategic Reading 3. Peer-Assisted Learning	Business as usual	Reading comprehension and vocabulary recognition and recall	CSR 26 SRE 27 PAL 25 C 30
Mohamadi (2018)	Iran	one group in Projectbased learning and another in Electronic Project-based Learning	business as usual	Idiom knowledge	30/30/30
Tian & Suppasetseree (2013)	Chinese Second year undergraduates	Online TBLT	Business as usual	Listening	46/46
Rajabi & Hashemiana (2015)	Iranian adolescents from different language institutes	Task-based instruction	Blended learning with the same material	Resumptive pronoun in Writing	20/20

Tutors' perceptions of the role of written feedback in promoting self-regulated learning in students: A case study of Durham

Suleyman Yildirim

School of Education, Durham University, Durham, UK

suleyman.yildirim@durham.ac.uk

Suleyman Yildirim is a PhD candidate in the School of Education at Durham University, researching the effect of assessment on student self-regulated learning in higher education.

Tutors' perceptions of the role of written feedback in promoting self-regulated learning in students: A case study of Durham

Students’ self-regulated learning is very important in the higher education as students are expected to construct their own knowledge. However, tutors are also supposed to help their students to overcome their difficulties when students study independently. One of the ways to help students is to provide written feedback, which is a crucial part of formative assessment, because in that case, tutors may indicate their students’ weaknesses and give advice to them about how to strengthen those weak learning points. Therefore, it can be said that written feedback is a useful tool to enable students to improve their learning. In the literature, while there is a lot of research regarding the effect of written feedback on student self-regulated learning by measuring students’ abilities, there are few studies about how tutors try to promote their students’ self-regulated learning. Therefore, this study aims to investigate how tutors use written feedback in order to contribute to their students’ self-regulated learning ability. In this study, 37 academics were interviewed from different departments at University of Durham. Phenomenography has been used as a research method to analyse the data.

Keywords: Self-regulated learning, Formative assessment, Written feedback, Phenomenography

1. Introduction

The number of people entering universities has been enormously increasing at the present time (Boud and Molloy, 2013). However, the amount of resources such as the number of academics and funding has not kept pace with this rise (Boud and Molloy, 2013). Thus, there is increased pressure on academic staff because they are responsible to improve their students' understanding. Some authors assume that student-centred learning might decrease this pressure on academics because in this learning method, teachers give responsibility to their students to help them to construct their own learning (Labuhn et al., 2010). Student-centred learning, described as students' becoming responsible for and actively involved in their own learning, is recognized to be more beneficial for the promotion of student learning and understanding than teacher-centred learning, defined as teachers' being responsible for their students' learning (Barzegar, 2012). Furthermore, Pintrich (2004) describes students as citizens of learning in student-centred learning environments and as tourists of learning in teacher-centred learning environments. So, some researchers argue that since students are able to construct knowledge for themselves in student-centred learning environments, each student's learning output and experience are unique (Duffy and Azevedo, 2015; DiBenedetto and Benbenutty, 2013). For this reason, student-centred learning pedagogy is more likely to lead students to improve their self-regulated learning abilities. Students who are able to use self-regulated learning strategies can reach desired learning goals and get deeply engaged in learning.

2. Literature Review

2.1. Self-Regulated Learning

In literature, self-regulated learning is described as students' taking responsibility to plan, monitor and control their own cognition, motivation and behaviour to reach their goals (Pintrich, 2004). To improve this skill, students should create goals, choose strategies to achieve these goals and observe their own developments towards the goals they set (Schunk, 1996). It is considered to be an important skill for learners as Zimmerman (2000) notes learners having self-regulated learning ability are more likely to learn more and have higher levels of academic satisfaction than those who do not have self-regulated learning ability.

Models of SRL consist of three different phases summarised below that are 'forethought and planning', 'performance monitoring' and 'reflection on performance' (Zimmerman, 2002). In the 'forethought and planning' phase, students analyse the learning objectives of a particular task and create a plan to reach the goals. In the 'performance monitoring' phase, students implement strategies to improve their learning and monitor the effectiveness of strategies they used whether their learning develops as well as they have adequate motivation to complete the task. In the last 'reflection on performance' phase, students assess their performance and outcomes to assess whether the strategies they selected worked or not. Students also manage their emotions about the outcomes of their learning effort. Self-evaluation and –reflection will affect their future planning, strategies and goals by applying these three phases in order. In other words, this diagram works as a loop (Zimmerman, 2002). If learners see something is not going well, they can go back and reorganize their phases from

the beginning or they can go back to the phase which they have problem to fix it to continue their ways towards the goals they set.

Research shows that self-regulated learners spend more time to learn topics, willingly answer the questions, seek out help from their peers, teachers or additional sources to achieve the learning goals (Labuhn et al., 2010; Elstad and Turmo, 2010). Additionally, self-regulated learners create the most suitable learning environment for themselves to develop a deep understanding of the subject matter (Kolovelonis et al., 2011). Therefore, briefly it can be said that students who are able to use the diagram drawn below are more likely to accomplish the goal they set.

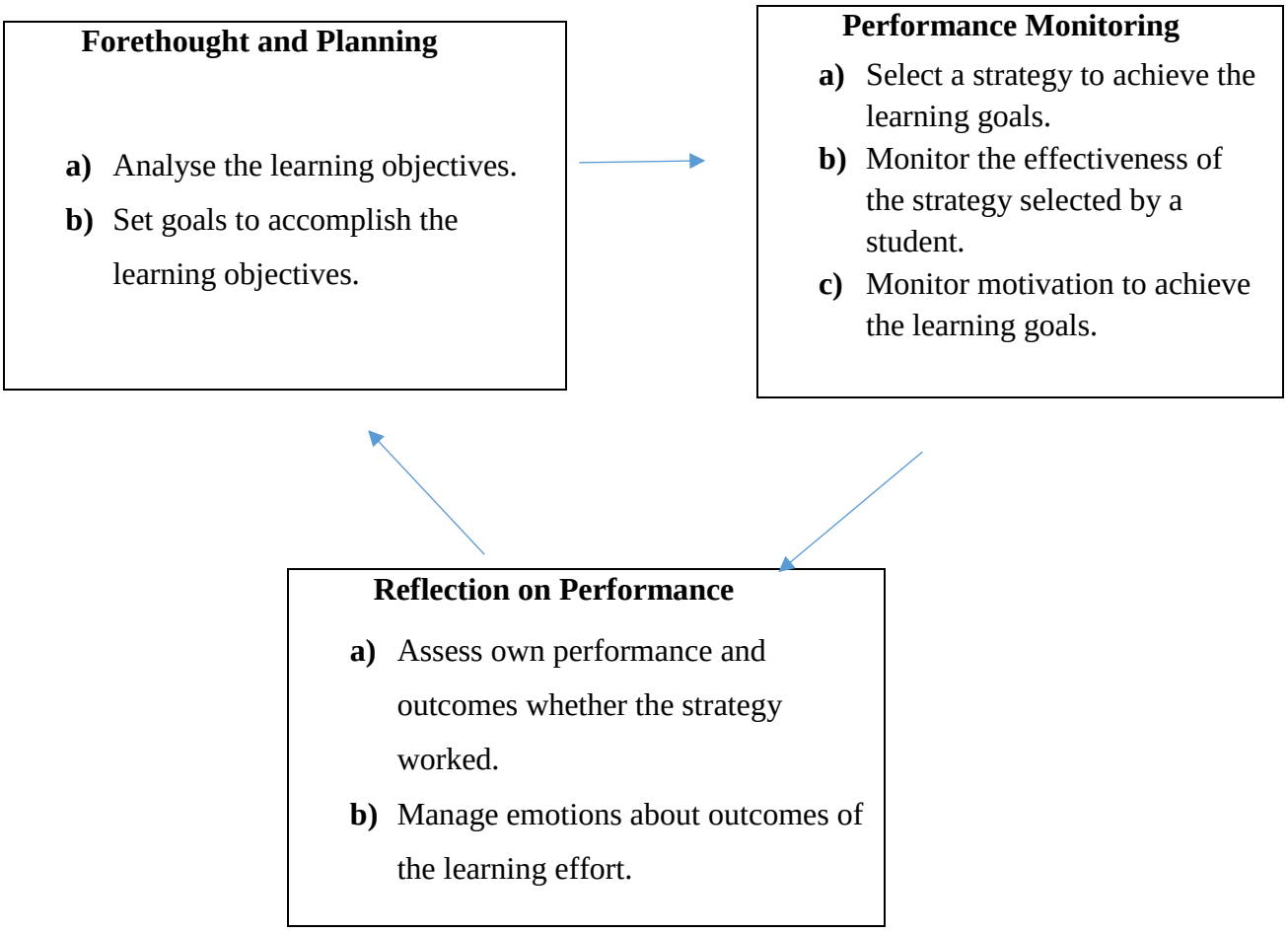


Figure 1. The phases of self-regulated learning based on Zimmerman (2002)

2.2. Formative Assessment

There are two well-known assessment models that are 'summative' and 'formative' (Black and Wiliam, 1998). Summative assessment is generally used to measure student performance in any subjects across the globe (Taras, 2005). Formative assessment is described as checking whether teaching instructions worked or not and tracking student progress rather than judging their understanding through exams or tests (Sadler, 1989). Formative assessment does not aim to judge students whether they are academically sufficient, it aims to determine students'

weaknesses and strengths to show them a way how to improve their weak skills in a particular domain of a subject (Steiner, 2016). Some researchers argue that applying merely summative assessment is not a good way to measure and improve student knowledge because it may lead students not to critically learn a task or a subject (Hattie and Timperley, 2007). Moreover, students may tend to memorize knowledge to receive high grade rather than further engage in a task (Sadler, 1989).

On the other hand, some academics argue that formative assessment enables students to become self-regulated learners because this assessment may stimulate student motivation that is described as the key factor of learning in higher education (Zimmerman, 2008; Pintrich, 2004). To determine whether formative assessment enhances student learning in the classroom, Black and Wiliam (1998) reviewed 250 journal articles and book chapters and then they conducted their own research. In their research, they measured students' test performances by comparing two groups that students who participated in formative activities and those who did not. According to the result, formative assessment is considerably helpful for student learning. Additionally, it increases all type of student group achievement, particularly low-achieving students too (Black and Wiliam, 1998a).

Formative assessment emphasizes the idea that feedback is a fundamental tool to contribute student learning because it provides information to students that they can use it to move forward (Brookhart, 2013). Furthermore, it raises students' awareness about their learning progress by indicating them the gap between their current academic skills and desired goals (Wiliam, 2011). Taras (2005) points out that to help students to close this gap on a task, detailed feedback that includes specific comments and advice how to correct their mistakes should be supplied to students as in this way students are more likely to pay attention to the task. Moreover, detailed feedback can be very effective if it manages to guide students to improve their weaknesses (Higgins, 2000; Shute, 2008). Hattie and Timperley (2007) report that detailed feedback can be given to students by using written feedback because teachers are able to express their opinion about their students' works in detail in that case.

2.3. Written Feedback

Written feedback is a good way to inform students about their mistakes and poor abilities and to advise them how to advance their poor abilities (Chong, 2018). Therefore, it may be said that written feedback enables students to take their own responsibilities to advance their knowledge, because teachers offer them some ways how to reach the learning goals (Xu, 2017). When students receive written feedback from their teachers, they are more likely to perceive better about the success criteria since students perceiving what they are supposed to do to achieve the desired learning goal may make sufficient effort to reach the goals (Carless, 2006). This also enables students to activate them as the owners of their own learning (Hattie, 2006).

Since students take their own responsibility to fill a gap between what is understood and what is aimed to be understood, they are more likely to accomplish their goal because students create their own studying style and learn how to learn by themselves (Steiner, 2016).

Furthermore, students can have adequate motivation as they have already begun to be able to discover the most suitable method for themselves to successfully attain their learning goals (Nicol and Macfarlane-Dick, 2006).

Teachers' advice given in written feedback to students how to reach the learning goals may also lead them to create a bridge between teachers and students as students can read what their teachers think about their work and what their weak abilities are and the way they should follow to achieve the learning goals (Boud and Molloy, 2013). Moreover, students have an opportunity to read their feedback when they need to remember what they are supposed to do to improve their skills (Xu, 2017). However, researchers emphasize that students sometimes do not perceive what teachers mean in their written feedback (Chong, 2018). Furthermore, students may define their written feedback as useless to improve their learning, while teachers think their feedback is considerably helpful for their students (Basturkmen et al., 2014). For this reason, this study aims to investigate how academics describe their written feedback and how they use it to improve students' learning and their self-regulated learning ability.

3. Methodology

This research focuses on the relation between written feedback and student self-regulated learning in higher education. Although most research has discussed the effect of written feedback on student academic progress in terms of students' perceptions and opinions, there is very little research about this topic investigating academics' perspectives of it. Therefore, I am planning to compare academics' perceptions and implementations about how they use their written feedback to help their students to become self-regulated learners. I also aim to investigate what kind of differences there are among various departments. So, in this research project, 37 academics (Assistant, Associate and Full Professor) have been interviewed in the Faculty of Arts and Humanities, the Faculty of Science and the Faculty of Social Sciences and Health at Durham University. Since I will analyse individuals' experiences about learning and teaching in the higher education, phenomenography seemed to be appropriate approach to answer my research questions:

1) How do academics conceptualise self-regulated learning?

2) What are academics' perceptions about how their written feedback enables their students to become self-regulated learners at Durham University?

3.1. Phenomenography

Phenomenography was formed and primarily developed as a new research approach by Swedish educational researchers during the 1970s. (Marton et al., 1977; Marton and Svensson, 1979; Säljö, 1979). Ference Marton firstly used phenomenography to investigate variation in student learning outcomes. The purpose of this research approach is to find out questions related to how people learn and understand knowledge in a specific context (Marton and Booth, 1997; Svensson, 1997). Using phenomenography research in different contexts to explore experience of learning leads it to include the most typical experiences

(Edwards, 2007). Moreover, the development of phenomenography has been still going on in the discipline of education (Marton and Booth, 1997; Svensson, 1997).

Phenomenography is most frequently defined as "a research method for mapping the qualitatively different ways in which people experience, conceptualise, perceive, and understand various aspects of, and various phenomena in, the world around them" (Marton, 1986, p. 31). Phenomenography is describing things of appearance and things which people have experienced because people collectively experience and understand phenomena in qualitatively different ways which are interrelated (Marton, 1986). In the phenomenography, the object and the research subjects are viewed together so it can be said that phenomenography is a relational approach to explore their relations (Limberg, 2000). So, when phenomenography creates inseparable relations between subject and object, the phenomenon as a whole is represented in that case. Individuals' experiences are very important for my research as I will try to analyse academics' perceptions and experiences about learning, teaching and assessment. I will also aim to discuss self-regulated learning construct under those educational tenets.

4. Discussion and Conclusion

Students should be able to use cognitive strategies to get engaged in their learning tasks (Xu, 2017). After using cognitive strategies, they need to use metacognitive strategies to monitor their own progress whether cognitive strategies they used worked or not (Zimmerman, 2008). Students' evaluations of their own learning progress are very important to organise their resources to benefit from them effectively and efficiently. In that case, students are able to regulate their own learning to attain the goals they set. To be able to use cognitive and metacognitive strategies to organise their own learning progress, students should have sufficient motivation because motivation energizes them to make effort to achieve their own learning goals (Wiliam, 2011).

If students describe that their learning tasks are important and students have real interest in them, they are more likely to be willing to complete their tasks. Students who are willing to improve their learning can set more challenging goals for themselves. If they achieve the goals they set, they will likely to improve their confidence too. Students' having sufficient confidence might set mastery goals as a next step (DiBenedetto and Bembenutty, 2013). In that way, students might get deeper understanding of a subject matter with time. As a result, students can gain self-regulated learning abilities so they can transfer those abilities to understand other subject matters (DiBenedetto and Bembenutty, 2013). Therefore, students are also able to create their own learning environment to meet their own needs to attain the goals they set.

Written feedback supports students to advance their weak skills and understanding by showing them what they need to do to improve those weaknesses (Chong, 2018). Students receiving written feedback from their teachers can set more challenging goals to push themselves to develop their learning abilities more. Therefore, it can be said that written feedback is one of student learning environment components that enables students to study independently to reach highest academic capabilities (Xu, 2017).

References

Barzegar, M. (2012). *The relationship between goal orientation and academic achievement. The mediation role of self-regulated learning strategies- a path analysis*. International Conference on Management, Humanity and Economics (ICMHE'2012) August 11-12, 2012 Phuket (Thailand).

Basturkmen, H., East, M., and Bitchener, J. (2014). "Supervisors' on-script feedback comments on drafts of dissertations: socialising students into the academic discourse community." *Teaching in Higher Education*, 19 (4): 432–445.

Black, P. & Wiliam, D. (1998). 'Assessment and classroom learning', *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.

Black, P., & Wiliam, D. (1998a). 'Inside the black box: raising standards through classroom assessment', *Phi Delta Kappan, 92*(1), 81–90.

Boud, D., and Molloy, E. (2013). "Rethinking Models of Feedback for Learning: The Challenge of Design." *Assessment & Evaluation in Higher Education*, 38 (6): 698-712.

Carless, D. (2006). "Differing Perceptions in the Feedback Process." *Studies in Higher Education* 31 (2): 219–233.

Chong, I. (2018). Interplay among technical, socio-emotional and personal factors in written feedback research, *Assessment & Evaluation in Higher Education*, 43:2, 185-196. DOI: 10.1080/02602938.2017.1317712

DiBenedetto, M. K., & Bembenutty, H. (2013). Within the pipeline: self-regulated learning, self-efficacy, and socialization among college students in science courses. *Learning & Individual Differences*, 23, 218-224.

Duffy, M., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338-348.

Elstad, E., & Turmo, A. (2010). Students' self-regulation and teacher's influence in science: Interplay between ethnicity and gender. *Research in Science & Technological Education*, 28 (3), 249-260.

Hattie, J. (2006). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Hattie, J. & Timperley, H. (2007). 'The power of feedback', *Review of Educational Research*, 77(1), 81-112.

Higgins, R. (2000). *'Be more critical!' Rethinking assessment feedback*. Paper presented at the BERA Conference Cardiff University, September 7–10.

Kolovelonis, A., Goudas, M., & Dermitzaki, I. (2011). The effect of different goals and self-recording on self-regulation of learning a motor skill in a physical education setting. *Learning and Instruction*, 21 (3), 355-364.

Labuhn, A.S., Zimmerman, B.J., & Hasselhorn, M. (2010). Enhancing students' self-regulation and mathematics performance: The influence of feedback and self-evaluative standards. *Metacognition and Learning*, 5 (2), 173-194.

Limberg, L. (2000). Phenomenography: a relational approach to research on information needs, seeking and use. *The New Review of Information Behaviour Research*, 1: 51-67.

Marton, F., Dahlgren, L. O., Svensson, L., & Säljö, R. (1977). *Learning and conception of the world around us*. Stockholm: Almquist and Wiksell.

Marton, F., & Svensson, L. (1979). Conceptions of research in student learning. *Higher Education*, 8: 471-486.

Marton, F. (1986). Phenomenography – a research approach to investigating different understandings of reality. *Journal of Thought*, 21(3): 28-49.

Marton, F. & Booth, S. (1997). *Learning and awareness*. Mahwah, NJ: L. Erlbaum Associates.

Nicol, D. & Macfarlane-Dick, D. (2006) 'Formative assessment and self-regulated learning: a model and seven principles of good feedback practice', *Studies in Higher Education, 31*(2), 199–218. doi: 10.1080/03075070600572090.

Pintrich, P.R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385–407. Retrieved from http://dx.doi.org/10.1007/s10648-004-0006-x.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.

Säljö, R. (1979). Learning about learning. *Higher Education*, 8, 443-451.

Schunk, D. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*(2), 359–382.

Shute, V. (2008) 'Focus on formative feedback', *ETS Research Report Series, 2007*(1), i–47. Retrieved from http://dx.doi.org/10.1002/j.2333-8504.2007.tb02053.x

Steiner, H. H. (2016). The strategy project: promoting self-regulated learning through an authentic assignment. *International Journal of Teaching and Learning in Higher Education, 2016*, Volume 28, Number 2, 271-282. Retrieved from http://www.isetl.org/ijtlhe/

Svensson, L. (1997). Theoretical foundations of phenomenography. *Higher Education Research & Development*, 16(2), 159-171.

Taras, M. (2005). 'Assessment – summative and formative – some theoretical reflections', *British Journal of Educational Studies, 53*(4), 466–478.

Xu, L. (2017). Written feedback in intercultural doctoral supervision: a case study, *Teaching in Higher Education*, 22:2, 239-255, DOI: 10.1080/13562517.2016.1237483

Wiliam, D. (2011). *Embedded Formative Assessment*. Bloomington, In: Solution Tree Press.

Zimmerman, B. (2000). Attaining self-regulated learning: a social-cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory into practice*, 41, 64–70. Retrieved from https://doi.org/10.1207/s15430421tip4102_2.

Zimmerman, B.J. (2008). Investigating self-regulation and motivation: historical background,

methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183.